

# **СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ**

**Научный журнал Российской академии наук  
(издается под руководством Отделения нанотехнологий  
и информационных технологий РАН)**

Издается с 1989 года  
Журнал выходит ежеквартально

**Учредители:**  
**Российская академия наук**  
**Институт проблем информатики Российской академии наук**

## **РЕДАКЦИОННЫЙ СОВЕТ**

академик РАН И. А. Соколов — председатель Редакционного совета  
академик РАН Г. И. Савин  
академик РАН А. Л. Стемпковский  
член-корреспондент РАН Ю. Б. Зубарев  
профессор Ш. Долев (S. Dolev, Beer-Sheva, Israel)  
профессор Ю. Кабанов (Yu. Kabanov, Besancon, France)  
профессор М. Никулин (M. Nikulin, Bordeaux, France)  
профессор В. Ротарь (V. Rotar, San-Diego, USA)  
профессор И. Ушаков (I. Ushakov, San-Diego, USA)  
профессор М. Финкельштейн (M. Finkelstein, Rostok, Germany)

## **РЕДАКЦИОННАЯ КОЛЛЕГИЯ**

академик РАН И. А. Соколов — главный редактор  
профессор, д.ф.-м.н. С. Я. Шоргин — заместитель главного редактора  
д.т.н. В. Н. Захаров  
проф., д.т.н. В. Д. Ильин проф., д.ф.-м.н. А. В. Печинкин  
проф., д.ф.-м.н. Л. А. Калиниченко проф., д.г.-м.н. Р. Б. Сейфуль-Мулюков  
д.т.н. В. А. Козмидиади проф., д.т.н. И. Н. Синицын  
проф., д.т.н. К. К. Колин к.т.н. А. В. Филин  
проф., д.ф.-м.н. В. Ю. Королев к.ф.-м.н. С. А. Христочевский

## **Редакция**

профессор, д.г.-м.н. Р. Б. Сейфуль-Мулюков  
к.ф.-м.н. Е. Н. Арутюнов  
С. Н. Стригина (ответственный секретарь)

© Институт проблем информатики Российской академии наук, 2014

Журнал входит в систему Российского индекса научного цитирования (РИНЦ):

**[http://elibrary.ru/title\\_about.asp?id=28980](http://elibrary.ru/title_about.asp?id=28980)**

Журнал включен в базу данных CrossRef (систему DOI — Digital Object Identifier),  
в базу данных Ulrich's periodicals directory  
и в информационную систему «Общероссийский математический портал Math-Net.Ru»

Журнал реферируется в «Реферативном журнале» ВИНТИ  
и в системе Google Scholar

Журнал «Системы и средства информатики»  
включен в «Перечень российских рецензируемых журналов,  
в которых должны быть опубликованы основные научные результаты диссертаций  
на соискание ученых степеней доктора и кандидата наук», утвержденный ВАК

**<http://www.ipiran.ru/journal/collected>**

# **СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ**

Том 24 № 3 Год 2014

## **СОДЕРЖАНИЕ**

Математическое обеспечение аналитического моделирования стохастических систем со сложными нелинейностями <i>И. Н. Синицын, В. И. Синицын, И. В. Сергеев, Э. Р. Корепанов, В. В. Белоусов, В. С. Шоргин</i>	<b>4</b>
Безопасные архитектуры распределенных систем <i>А. А. Грушо, Н. А. Грушо, Е. Е. Тимонина, С. Я. Шоргин</i>	<b>18</b>
Скрытые каналы в беспроводных сетях стандарта 802.11 <i>Н. А. Грушо</i>	<b>32</b>
Умножитель с накоплением: методологические аспекты <i>И. А. Соколов, Ю. А. Степченков, С. Г. Бобков, Ю. В. Рождественский, Ю. Г. Дьяченко</i>	<b>44</b>
Самосинхронный умножитель с накоплением: практическая реализация <i>Ю. А. Степченков, Ю. Г. Дьяченко, Ю. В. Рождественский, Н. В. Морозов, Д. Ю. Степченков, А. В. Рождественскене, А. В. Сурков</i>	<b>63</b>
Современные тенденции в развитии архитектур интегрированных сетевых процессоров <i>В. Б. Егоров</i>	<b>78</b>
Двухпараметрический анализ магнитно-резонансного изображения методом максимума правдоподобия в сравнении с однопараметрическим приближением <i>Т. В. Яковлева, Н. С. Кульберг</i>	<b>92</b>
Свойства оконной дисперсии миограммы как случайного процесса <i>М. Ш. Хазиахметов</i>	<b>110</b>

# **СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ**

Том 24 № 3 Год 2014

## **СОДЕРЖАНИЕ**

Исследование лабораторного прототипа искусственной гетерогенной системы для диагностики артериальной гипертензии

**И. А. Кириков, А. В. Колесников, С. Б. Румовская 121**

Особенности формирования номенклатуры и количества компонентов ЗИП в автоматизированных информационных системах в защищенном исполнении

**А. А. Зацаринный, С. В. Козлов, А. И. Гаранин 144**

Особенности разработки документации для автоматизированных информационных систем в защищенном исполнении

**А. А. Зацаринный, Ю. С. Ионенков 156**

Мониторинг информационной безопасности как облачный сервис

**О. А. Бондаренко, К. И. Волович, В. А. Кондрашев 169**

Автоматизированная информационная система комплексного мониторинга телекоммуникационной сети

**А. А. Сорокин, А. Г. Тарасов, С. П. Королёв 176**

Систематизация знаний о программируемых задачах

**А. В. Ильин, В. Д. Ильин 192**

Использование хеш-функций для повышения скорости морфологического анализа русских текстов

**Н. В. Сомин, М. М. Шарнин 204**

К вопросу о логистике внедрения и использования мобильных электронных учебников

**Д. А. Богданова, А. А. Федосеев 218**

Об авторах **230**

## МАТЕМАТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ АНАЛИТИЧЕСКОГО МОДЕЛИРОВАНИЯ СТОХАСТИЧЕСКИХ СИСТЕМ СО СЛОЖНЫМИ НЕЛИНЕЙНОСТЯМИ\*

*И. Н. Синицын<sup>1</sup>, В. И. Синицын<sup>2</sup>, И. В. Сергеев<sup>3</sup>, Э. Р. Корепанов<sup>4</sup>,  
В. В. Белоусов<sup>5</sup>, В. С. Шоргин<sup>6</sup>*

**Аннотация:** Изучаются конечномерные дифференциальные стохастические системы (ДСтС) со сложными конечными, дифференциальными и интегральными нелинейностями, с винеровскими и пуассоновскими шумами, а также эредитарные (интегродифференциальные) стохастические системы (ЭСтС), приводимые к ДСтС. Представлены уравнения методов нормальной аппроксимации (МНА) и статистической линеаризации (МСЛ) для аналитического моделирования нестационарных и стационарных нормальных (гауссовских) процессов в нелинейных ДСтС и нелинейных ЭСтС, полученных путем аппроксимации эредитарных ядер для дифференцируемых нелинейностей линейными операторными уравнениями и для недифференцируемых нелинейностей сингулярными ядрами. Рассматриваются методы вычисления типовых интегралов МНА (МСЛ) для сложных (многомерных и векторного аргумента) конечных, дифференциальных и интегральных нелинейностей и алгоритмы интегрирования уравнений МНА (МСЛ). Разработано инструментальное программное обеспечение в среде MATLAB в составе комплекса «IDStS», а также тестовые примеры.

**Ключевые слова:** аналитическое моделирование; дифференциальные стохастические системы с винеровскими и пуассоновскими шумами; инструментальное программное обеспечение «IDStS»; математическое обеспечение; метод нормальной аппроксимации; метод статистической линеаризации; сложные иррациональные нелинейности; сложные конечные, дифференциальные и интегральные нелинейности; эредитарные стохастические системы, приводимые к дифференциальному

**DOI:** 10.14357/08696527140301

---

\*Работа выполнена при частичной финансовой поддержке программы «Интеллектуальные информационные технологии, системный анализ и автоматизация» (проект 1.7).

<sup>1</sup>Институт проблем информатики Российской академии наук, sinitsin@dol.ru

<sup>2</sup>Институт проблем информатики Российской академии наук, vsinitsin@ipiran.ru

<sup>3</sup>Институт проблем информатики Российской академии наук, isergeev@ipiran.ru

<sup>4</sup>Институт проблем информатики Российской академии наук, ekorepanov@ipiran.ru

<sup>5</sup>Институт проблем информатики Российской академии наук, vbelousov@ipiran.ru

<sup>6</sup>Институт проблем информатики Российской академии наук, vshorgin@ipiran.ru

## 1 Введение

Как известно [1, 2], математическими моделями функционирования многих современных технических систем информатики служат стохастические системы (СтС), описываемые дифференциальными, интегральными и интегродифференциальными уравнениями со сложными дробно-рациональными, иррациональными и интегральными нелинейностями. В [1] дано систематическое изложение МНА и МСЛ для ДСтС и ЭСтС, приводимых к дифференциальным. В [2] рассмотрены МНА и МСЛ для аналитического моделирования нормальных стохастических процессов (СтП) на случай СтС со сложными конечными, дифференциальными и интегральными нелинейностями.

Статья посвящена математическому обеспечению для аналитического моделирования ДСтС, а также ЭСтС, приводимых к ДСтС со сложными конечными, дифференциальными и интегральными нелинейностями, на основе МНА и МСЛ, разработанных в [2].

## 2 Уравнения методов нормальной аппроксимации и статистической линеаризации для дифференциальных стохастических систем и эредитарных стохастических систем, приводимых к дифференциальным

Как известно [3–5], уравнения конечномерных непрерывных нелинейных систем со стохастическими возмущениями путем введения инструментальных переменных и расширения вектора состояния ДСтС могут быть записаны в виде следующего векторного стохастического дифференциального уравнения Ито:

$$dY_t = a(Y_t, \Theta, t) dt + b(Y_t, \Theta, t) dW_0 + \int_{R_0^q} c(Y_t, \Theta, t, v) P^0(dt, dv), \\ Y(t_0) = Y_0. \quad (1)$$

Здесь  $Y_t$  —  $p$ -мерный вектор состояния ДСтС;  $\Theta$  —  $p_\Theta$ -мерный вектор случайных параметров, в общем случае зависящих от  $t$ ;  $a = a(Y_t, \Theta, t)$  и  $b = b(Y_t, \Theta, t)$  — известные  $(p \times 1)$ -мерная и  $(p \times r)$ -мерная функции  $Y_t$  и  $t$ ;  $W_0 = W_0(t)$  —  $r$ -мерный винеровский СтП интенсивности  $\nu_0 = \nu_0(t)$ ;  $c(Y_t, \Theta, t, v)$  —  $(p \times 1)$ -мерная функция  $Y_t, t$  и вспомогательного  $(q \times 1)$ -мерного параметра  $v$ ;  $\int dP^0(t, A)$  — центрированная пуассоновская мера, определяемая следующим образом:

$$\int_{\Delta} dP^0(t, A) = \int_{\Delta} dP(t, A) = \int_{\Delta} \nu_P(t, A) dt,$$

где  $\int\limits_{\Delta} dP(t, A)$  — число скачков пуассоновского СтП в интервале времени  $\Delta = [t_1, t_2]$ ;  $\nu_P(t, A)$  — интенсивность пуассоновского СтП  $P(t, A)$ ;  $A$  — некоторое борелевское множество пространства  $R_0^q$  с выколотым началом. Начальное значение  $Y_0$  представляет собой случайную величину (СВ), не зависящую от приращений СтП  $W_0(t)$  и  $P(t, A)$  на интервалах времени, следующих за  $t_0$ ,  $t_0 \leq t_1 \leq t_2$ , для любого множества  $A$ .

В случае аддитивных нормальных (гауссовских) и пуассоновских возмущений уравнение (1) имеет вид:

$$\dot{Y}_t = a(Y_t, \Theta, t) + b_0(\Theta, t)V, \quad V = \dot{W}, \quad Y(t_0) = Y_0. \quad (2)$$

Здесь  $W$  — СтП с независимыми приращениями, представляющий собой смесь нормального и обобщенного пуассоновского СтП.

В некоторых случаях вводят расширенный вектор состояния  $\bar{Y}_t = [Y_t^T \Theta^T]^T$ . Тогда уравнение ДСтС (1) совместно с уравнениями формирующего фильтра для  $\Theta$  приводятся к (1), если положить

$$\left. \begin{aligned} \bar{a} &= \bar{a}(\bar{Y}_t, t) = [a(Y_t, \Theta, t)^T a^\Theta(\Theta, t)^T]^T; \\ \bar{b} &= \bar{b}(\bar{Y}_t, t) = [b(Y_t, \Theta, t)^T b^\Theta(\Theta, t)^T]^T; \\ \bar{c} &= \bar{c}(\bar{Y}_t, t) = [c(Y_t, \Theta, t, v)^T c^\Theta(\Theta, t)^T]^T. \end{aligned} \right\} \quad (3)$$

В дальнейшем в общем случае будем рассматривать ДСтС (1) или (2) при условии (3) и опускать черту над  $\bar{Y}_t$  и  $\bar{a}, \bar{b}, \bar{c}$ .

Если предположить существование конечных вероятностных моментов второго порядка для моментов  $t_1$  и  $t_2$ , то уравнения МНА примут следующий вид [3, 4]:

- для характеристических функций

$$g_1^N(\lambda; t) = \exp \left[ i\lambda^T m_t - \frac{1}{2} \lambda^T K_t \lambda \right]; \quad (4)$$

$$g_{t_1, t_2}^N(\lambda_1, \lambda_2; t_1, t_2) = \exp \left[ i\bar{\lambda}^T \bar{m}_2 - \frac{1}{2} \bar{\lambda}^T \bar{K}_2 \bar{\lambda} \right], \quad (5)$$

где

$$\bar{\lambda} = [\lambda_1^T \lambda_2^T]^T; \quad \bar{m}_2 = [m_{t_1}^T m_{t_2}^T]^T; \quad \bar{K}_2 = \begin{bmatrix} K(t_1, t_1) & K(t_1, t_2) \\ K(t_2, t_1) & K(t_2, t_2) \end{bmatrix};$$

- для математических ожиданий  $m_t$ , ковариационной матрицы  $K_t$  и матрицы ковариационных функций  $K(t_1, t_2)$ :

$$\dot{m}_t = a_1(m_t, K_t, t), \quad m_0 = m(t_0); \quad (6)$$

$$\dot{K}_t = a_2(m_t, K_t, t), \quad K_0 = K(t_0); \quad (7)$$

$$\frac{\partial K(t_1, t_2)}{\partial t_2} = K(t_1, t_2) a_{21}(m_{t_2}, K_{t_2}, t_2)^T, \quad K(t_1, t_1) = K_{t_1}. \quad (8)$$

Здесь приняты следующие обозначения:

$$a_1 = a_1(m_t, K_t, t) = \mathbf{M}_N a(Y_t, t); \quad (9)$$

$$a_2 = a_2(m_t, K_t, t) = a_{21}(m_t, K_t, t) + a_{21}(m_t, K_t, t)^T + a_{22}(m_t, K_t, t); \quad (10)$$

$$a_{21} = a_{21}(m_t, K_t, t) = \mathbf{M}_N a(Y_t, t) Y_t^{0T}; \quad (11)$$

$$a_{22} = a_{22}(m_t, K_t, t) = \mathbf{M}_N \sigma(Y_t, t); \quad (12)$$

$$\sigma(Y_t, t) = b(Y_t, t) \nu_0(t) b(Y_t, t)^T + \int_{R_0^q} c(Y_t, t, v) c(Y_t, t, v)^T \nu_P(t, dv); \quad (13)$$

$$m_t = \mathbf{M} Y_t, \quad Y_t^0 = Y_t - m_t; \quad K_t = \mathbf{M}_N Y_t^0 Y_t^{0T}; \quad K(t_1, t_2) = \mathbf{M}_N Y_{t_1}^0 Y_{t_2}^0,$$

где  $\mathbf{M}_N$  — символ вычисления математического ожидания для нормальных распределений (4) и (5).

Если существуют нормальные стационарные СтП вида  $m_t = m^*$ ,  $K_t = K^*$ ,  $K(t_1, t_2) = k(\tau)$  ( $\tau = t_1 - t_2$ ), то они определяются уравнениями [3, 4]:

$$a_1(m^*, K^*) = 0; \quad a_2(m^*, K^*) = 0; \quad (14)$$

$$\left. \begin{aligned} \dot{k}_\tau(\tau) &= a_{21}(m^*, K^*) K^{*-1} k(\tau), \quad k(0) = K^* \quad (\forall \tau > 0); \\ k(\tau) &= k(-\tau)^T \quad (\forall \tau < 0). \end{aligned} \right\} \quad (15)$$

При этом необходимо, чтобы матрица  $a_{21}(m^*, K^*)$  была асимптотически устойчивой.

Для ДСтС (2) уравнения (4)–(8) переходят в уравнения МСЛ Казакова [3, 4], если принять

$$a(Y_t, t) = a_1(m_t, K_t) + k_1^a(m_t, K_t) Y_t^0; \quad (16)$$

$$b(Y_t, t) = b_0(t); \quad \sigma(Y_t, t) = b_0(t) \nu(t) b_0(t)^T = \sigma_0(t); \quad (17)$$

$$k_1^a(m_t, K_t, t) = \left[ \left( \frac{\partial}{\partial m_t} \right) a_0(m_t, K_t, t)^T \right]^T; \quad (18)$$

$$\dot{m}_t = a_1(m_t, K_t, t), \quad m_0 = m(t_0); \quad (19)$$

$$\dot{K}_t = k_1^a(m_t, K_t, t)K_t + K_t k_1^a(m_t, K_t, t)^T + \sigma_0(t), \quad K_0 = K(t_0); \quad (20)$$

$$\frac{\partial K(t_1, t_2)}{\partial t_2} = K(t_1, t_2)K_{t_2}k_1^a(m_{t_2}, K_{t_2}, t_2)^T, \quad K(t_1, t_1) = K_{t_1}. \quad (21)$$

В основе МСЛ для стационарных ДСтС (2) при условии асимптотической устойчивости матрицы  $k_1^a(m^*, K^*)$  лежат уравнения (14) и (15), записанные в виде:

$$a_1(m^*, K^*) = 0; \quad (22)$$

$$k_1^a(m^*, K^*)K^* + K^*k_1^a(m^*, K^*)^T + \sigma_0^* = 0; \quad (23)$$

$$\left. \begin{aligned} \dot{k}_\tau(\tau) &= k_1^a(m^*, K^*)k(\tau), \quad k(0) = K^* \quad (\forall \tau > 0); \\ k(\tau) &= k(-\tau)^T \quad (\forall \tau < 0). \end{aligned} \right\} \quad (24)$$

Таким образом, уравнения (4)–(8) лежат в основе МНА для ДСтС (1), а (16)–(21) — в основе МСЛ для ДСтС (2). Для определения стационарных СтП в стационарных ДСтС согласно МНА служат соотношения (14), (15), а МСЛ — (22)–(24).

Теперь рассмотрим ЭСтС, описываемую интегродифференциальным уравнением Ито следующего вида [2, 5]:

$$\begin{aligned} dX_t = & \left[ a(X_t, \Theta, t) + \int_{t_0}^t a_1(X(\tau), \Theta, \tau, t) d\tau \right] dt + \\ & + \left[ b(X_t, \Theta, t) + \int_{t_0}^t b_1(X(\tau), \Theta, \tau, t) d\tau \right] dW_0 + \\ & + \int_{R_0^q} \left[ c(X_t, \Theta, t, v) + \int_{t_0}^t c_1(X(\tau), \Theta, \tau, t, v) d\tau \right] dP^0(t, dv) \quad (25) \end{aligned}$$

с начальным условием  $X(t_0) = X_0$ . В (25) сохранены обозначения (1). Функции  $a = a(X_t, \Theta, t)$ ,  $a_1 = a_1(X(\tau), \Theta, \tau, t)$ ,  $b = b(X_t, \Theta, t)$ ,  $b_1 = b_1(X(\tau), \Theta, \tau, t)$ ,  $c = c(X_t, \Theta, t, v)$  и  $c_1 = c_1(X(\tau), \Theta, \tau, t, v)$  имеют соответственно размерности  $p \times 1$ ,  $p \times 1$ ,  $p \times r$ ,  $p \times r$ ,  $p \times 1$  и  $p \times 1$  и допускают представления следующего вида:

$$\left. \begin{aligned} a_1 &= A(t, \tau)\varphi(X(\tau), \Theta, \tau); \\ b_1 &= B(t, \tau)\psi(X(\tau), \Theta, \tau); \\ c_1 &= C(t, \tau)\chi(X(\tau), \Theta, \tau, v). \end{aligned} \right\} \quad (26)$$

Здесь эредитарные ядра  $A = A(t, \tau) = [A_{ij}(t, \tau)]$  ( $i, j = \overline{1, p}$ ),  $B = B(t, \tau) = [B_{il}(t, \tau)]$  ( $i = \overline{1, p}, l = \overline{1, r}$ ) и  $C = C(t, \tau) = [C_{ij}(t, \tau)]$  ( $i, j = \overline{1, p}$ ) имеют соответственно размерности  $p \times p$ ,  $p \times r$  и  $p \times p$ . Они удовлетворяют следующим условиям физической реализуемости и асимптотического затухания:

$$A_{ij}(t, \tau) = 0; \quad B_{il}(t, \tau) = 0; \quad C_{ij}(t, \tau) = 0 \quad \forall \tau > t;$$

$$\int_{-\infty}^{\infty} |A_{ij}(t, \tau)| d\tau < \infty; \quad \int_{-\infty}^{\infty} |B_{il}(t, \tau)| d\tau < \infty; \quad \int_{-\infty}^{\infty} |C_{ij}(t, \tau)| d\tau < \infty.$$

В дальнейшем ограничимся случаем, когда эредитарные ядра удовлетворяют линейным операторным уравнениям [2, 5]. Нелинейные в общем случае функции  $\varphi = \varphi(X(\tau), \Theta, \tau)$ ,  $\psi = \psi(X(\tau), \Theta, \tau)$  и  $\chi = \chi(X(\tau), \Theta, \tau, v)$  имеют размерности  $p \times 1$ ,  $p \times p$  и  $p \times 1$  соответственно.

Важный класс ядер представляют собой сингулярные (вырожденные) ядра, когда имеют место представления [2, 5]:

$$A_{ij}(t, \tau) = A_{ij}^+(t)A_{ij}^-(\tau); \quad B_{il}(t, \tau) = B_{il}^+(t)B_{il}^-(\tau); \quad C_{ij}(t, \tau) = C_{ij}^+(t)C_{ij}^-(\tau), \\ i, l = \overline{1, p}, \quad j = \overline{1, r}.$$

В [2, 5] показано, что для дифференцируемых нелинейных функций  $\varphi$ ,  $\psi$  и  $\chi$  в (26) путем расширения вектора состояния за счет инструментальных переменных, аппроксимируемых линейными операторными уравнениями, определяющими эредитарные ядра, ЭСтС (25) приводится к ДСтС вида (1) или (2). В случае недифференцируемых нелинейных функций  $\varphi$ ,  $\psi$  и  $\chi$  в (26) ЭСтС (25) приводится к (1) или (2) на основе аппроксимации вырожденными (сингулярными) ядрами [2, 5].

Таким образом, после приведения ЭСтС (25) к ДСтС (1) или (2) можно воспользоваться уравнениями МНА и МСЛ для ДСтС.

### 3 Математическое обеспечение аналитического моделирования процессов в дифференциальных стохастических системах со сложными нелинейностями

Как известно [2, 7], различают скалярные и векторные (скалярного и векторного аргумента) нелинейности

$$Z_t = \varphi(Y_t, t), \tag{27}$$

описываемые элементарными функциями (ЭФ), специальными функциями (СФ), а также разрывными функциями.

Следуя [8], назовем скалярные функции скалярного аргумента, значения которых получаются с помощью конечного числа вычислительных операций над аргументом, зависимой переменной и постоянными числами ЭФ. Под вычислительными операциями понимают четыре арифметических действия, возвведение в целую или дробную степень, взятие тригонометрических или обратных им функций, логарифмирование и потенцирование. Они непрерывны. Элементарные функции делятся на алгебраические (степенные, многочлены, дробно-рациональные функции, иррациональные функции, являющиеся суперпозицией степенных функций с нецелыми показателями и др.) и трансцендентные (показательные логарифмические, тригонометрические функции, обратные тригонометрические функции и др.). Класс ЭФ обладает той замечательной особенностью, что производная ЭФ всегда является ЭФ. При этом операция интегрирования может выводить из класса ЭФ.

Класс СФ значительно шире класса ЭФ. К СФ принято относить функции, встречающиеся при решении задач математической физики, теории вероятностей и математической статистики и др. [8].

Интегральными функциями (ИФ) называются СФ, выражющиеся через интегралы от простых ЭФ, а также от функций, получающихся из них с помощью конечного числа вычислительных операций и операций дифференцирования [8]. Сюда относятся ИФ, обратные упомянутым. Наиболее часто используемыми ИФ являются: гамма-функция, интеграл вероятности и обратная ему функция, интегральные тригонометрические, показательная и логарифмическая функции, функции  $\gamma$ - и  $\chi^2$ -распределения и др. Среди других СФ следует выделить часто встречающиеся цилиндрические функции (функции Бесселя), эллиптические функции, а также ортогональные многочлены [8].

Для качественных исследований часто рассматривают скалярные нелинейности, лежащие в первом и третьем секторах и их векторные аналоги [6, 7]. При этом выделяют подклассы непрерывных, дифференцируемых, а также убывающих и неубывающих нелинейностей.

К разрывным нелинейностям относятся кусочно-постоянные и кусочно-непрерывные функции, а также функции конечной вариации. Важным частным классом скалярных нелинейностей являются функции, описываемые функциями ограниченной вариации. Специального рассмотрения заслуживают нелинейности со множествами устранимого и неустранимого разрыва.

Упомянем также неоднозначные нелинейности с зонами нечувствительности, а также гистерезисного типа, описываемые конечными функциями вида [3, 4]

$$Z_t = \varphi(Y_t, \dot{Y}_t, \dots, Y_t^{(n)}, t) \quad (28)$$

или функционалами

$$Z_t = \varphi(Y_{t_0}^t, t, Y_{t_0}). \quad (29)$$

Важным частным классом нелинейностей (29) являются эредитарные нелинейности, описываемые интегральными уравнениями вида

$$Z_t = \int_{t_0}^t A(t, \tau) \varphi(Y_\tau, \tau) d\tau, \quad (30)$$

где ядро  $A(t, \tau)$  определено (26).

Наконец, отметим, что наряду с интегральными нелинейностями (30) рассматриваются операторные нелинейности вида [3, 4]

$$Z_s = \sum_{\rho=1}^R \mathcal{A}_\rho \varphi_\rho. \quad (31)$$

Здесь  $\mathcal{A}_\rho, \dots, \mathcal{A}_R$  — произвольные линейные операторы, действующие над функциями  $l$  переменных  $t_1, \dots, t_l; \varphi_\rho = \varphi_\rho(Y_{t_1}, \dots, Y_{t_l}, s)$ . Такие нелинейности называются приводимыми к линейным [3, 4]. В частности, примерами (31) могут служить интегральные нелинейности вида:

$$\begin{aligned} Z_s &= \int_T \varphi(Y_t, t, s) dt; \\ Z_s &= \int_T \cdots \int_T \varphi(Y_{t_1}, \dots, Y_{t_l}; t_1, \dots, t_l, s) dt_1 \cdots dt_l. \end{aligned}$$

Как следует из уравнений (9)–(13), для МНА необходимо уметь вычислять следующие интегралы:

$$I_0^q = I_0^a(m_t, K_t, t) = a_1(m_t, K_t, t) = \mathbf{M}_N a(Y_t, t); \quad (32)$$

$$I_1^a = I_1^a(m_t, K_t, t) = a_{21}(m_t, K_t, t) = \mathbf{M}_N a(Y_t, t) Y_t^{0T}; \quad (33)$$

$$I_0^{\bar{a}} = I_0^{\bar{a}}(m_t, K_t, t) = a_{22}(m_t, K_t, t) = \mathbf{M}_N \bar{a}(Y_t, t). \quad (34)$$

При этом для МСЛ достаточно вычислить интеграл (32), причем интеграл  $I_1^a$  вычисляется по формуле [3, 4]:

$$k_1^a = k_1^a(m_t, K_t, t) = \left[ \left( \frac{\partial}{\partial m_t} \right) I_0^a(m_t, K_t, t)^T \right]^T. \quad (35)$$

Для аналитического моделирования нормальных СтП в ДСтС со сложными нелинейностями в составе инструментального программного комплекса «IDStS» в среде MATLAB разработано специализированное программное обеспечение и комплекс тестовых примеров. Для конечных сложных аналитически заданных нелинейных функций  $a(Y, t)$  для вычисления интегралов  $I_0^a$  и  $I_0^{\bar{a}}$  использовались символьные методы полиномиальной аппроксимации (см. подробно [9, 10]).

Для сложных нелинейностей, выражаемых интегральными, дифференциальными и другими операторными уравнениями, предварительно использовались алгоритмы приведения уравнений к дифференциальным [2, 5, 11–13]. При этом для численного интегрирования уравнений МНА (МСЛ) использовались стандартные одношаговые методы для жестких систем дифференциальных уравнений [9].

В приложении приведены тестовые примеры.

#### 4 Заключение

На основе [2] разработаны символьные алгоритмы, а также комплекс тестовых примеров для аналитического моделирования нормальных процессов в ДСтС и ЭСтС, приводимых к ДСтС со сложными конечными, дифференциальными и интегральными нелинейностями. Даны результаты тестирования в составе инструментального программного комплекса в среде MATLAB «IDStS» [13]. Приведены тестовые примеры.

Результаты допускают обобщение на случай ДСтС и ЭСтС со стохастическими нелинейностями, заданными конечными, дифференциальными и интегральными соотношениями или каноническими разложениями и представлениями [1, 3, 4, 10].

#### Приложение

В [1] для типовых степенных, показательных, тригонометрических и кусочно-постоянных нелинейностей (27) и (28) скалярного и векторного аргумента приведены обширные таблицы формул для интегралов (32)–(35). Скалярные нелинейности были использованы в качестве тестовых примеров.

Дополнительно в состав «IDStS» включены формулы для логарифмических, обратных тригонометрических, гиперболических и обратных гиперболических функций. Особое внимание было удалено сложным иррациональным нелинейностям (ИРН).

**Пример 1.** Рассмотрим вычисление интегралов (32) и (35) для одномерных ИРН скалярного аргумента

$$\varphi(Y_t, t) = |Y_t|^{\alpha-1} \operatorname{sgn} Y_t, \quad (36)$$

где  $\alpha$  — нецелый показатель.

Пользуясь (15) и (18), представим (36) в виде:

$$\varphi(Y_t, t) = \varphi_0(m_t, D_t, t, \alpha) + k_1^\varphi(m_t, D_t, t, \alpha)Y_t^0.$$

Здесь введены следующие обозначения:

$$\left. \begin{aligned} \varphi_0(m_t, D_t, t; \alpha) &= \Gamma(\alpha)D_t^{1/2}e^{-\zeta^2/4}D_{-\alpha}(\zeta); \\ k_1^\varphi(m_t, D_t, t; \alpha) &= \frac{\partial \varphi_0(m_t, D_t, t; \alpha)}{\partial m_t}, \end{aligned} \right\} \quad (37)$$

где  $\Gamma(\alpha)$  — гамма-функция;  $\zeta = m_t/\sqrt{D_t}$  — отношение сигнал–шум;  $D_{-\alpha}(\zeta)$  — функция параболического цилиндра [14]. При вычислении были учтены следующие соотношения [8, 14]:

$$\int_0^\infty x^{\alpha-1} e^{-\beta x^2 - \gamma x} dx = (2\beta)^{-\alpha/2} \Gamma(\alpha) \exp\left(\frac{\gamma^2}{8\beta}\right) D_{-\alpha}\left(\frac{\gamma}{\sqrt{2\beta}}\right);$$

$$\begin{aligned} \frac{dD_\rho(\zeta)}{d\zeta} &= -\frac{\zeta}{2} D_\rho(\zeta) - \rho D_{\rho-1}(\zeta) = \\ &= \frac{\zeta}{2} D_\rho(\zeta) - D_{\rho+1}(\zeta) \quad (\operatorname{Re} \beta > 0, \operatorname{Re} \alpha > 0, \rho = -\alpha). \end{aligned}$$

Соотношения (37) могут быть использованы для вычисления интегралов (34).

Подробно рассмотрены сложные ИРН вида  $\varphi(Y_t, t) = \tilde{\varphi}(|Y_t|^{\alpha-1}, t, \alpha)$ , где функции  $\tilde{\varphi}$  являются типовыми кусочно-постоянными функциями [1].

**Пример 2.** Для нелинейной ДСтС вида

$$\dot{Y}_t = -a_t |Y_t|^{\alpha-1} \operatorname{sgn} Y_t + b_t V$$

( $\alpha, a_t$  и  $b_t$  — постоянные;  $V$  — белый шум интенсивности  $\nu_t$ ) МСЛ приводит к следующим уравнениям:

$$\begin{aligned} \dot{m}_t &= -a_t \varphi_0(m_t, D_t, t; \alpha); \\ \dot{D}_t &= -2a_t k_1^\varphi(m_t, D_t, t; \alpha) + b_t^2 \nu_t, \end{aligned}$$

где  $\varphi_0$  и  $k_1^\varphi$  определены в (37). Относительная точность при  $a_t = 1, b_t = 0, 1, \nu_t = 1$  и  $\alpha = 0,5$  составляет 6%.

**Пример 3.** Для ДСтС

$$\dot{Y}_t = -a_t Y_t + b_t (|Y_t|^{\delta-1} \operatorname{sgn} Y_t) V$$

с учетом примера 1 МНА дает следующие уравнения:

$$\begin{aligned} \dot{m}_t &= -a_t m_t; \\ \dot{D}_t &= -2a_t D_t + b_t^2 \nu_t \varphi_0(m_t, D_t, t, \delta). \end{aligned}$$

Здесь  $\varphi_0(m_t, D_t, t, \delta)$  определяется (37) при  $\delta = 2\alpha - 1$ . Относительная точность аналитического моделирования стационарного СтП при  $a_t = 1, b_t = 1, \nu_t = 10^{-1}$  и  $\alpha = 0,5$  составляет 15%.

## Литература

1. Синицын И. Н., Синицын В. И. Лекции по нормальной и эллипсоидальной аппроксимации распределений в стохастических системах. — М.: ТОРУС ПРЕСС, 2013. 488 с.

2. Синицын И. Н., Синицын В. И. Аналитическое моделирование нормальных процессов в стохастических системах со сложными нелинейностями // Информатика и её применения, 2014 (в печати). Т. 8. Вып. 3. doi: 10.14357/19922264140302.
3. Пугачев В. С., Синицын И. Н. Стохастические дифференциальные системы. Анализ и фильтрация. — М.: Наука, 1990. 632 с.
4. Пугачев В. С., Синицын И. Н. Теория стохастических систем. — М.: Логос, 2000; 2004. 1000 с.
5. Синицын И. Н. Анализ и моделирование распределений в эредитарных стохастических системах // Информатика и её применения, 2014. Т. 8. Вып. 1. С. 2–11.
6. Воронов А. А. Устойчивость, управляемость, наблюдаемость. — М.: Наука. 1979. 336 с.
7. Воронов А. А. Введение в динамику сложных управляемых систем. — М.: Наука, 1985. 352 с.
8. Попов Б. А., Теслер Г. С. Вычисление функций на ЭВМ: Справочник. — Киев: Наукова Думка, 1984. 599 с.
9. Корн Г., Корн Т. Справочник по математике / Пер. с англ. — М.: Наука, 1984. (Korn G. A., Korn T. M. Mathematical handbook. — New York: McGraw-Hill, 1968.)
10. Синицын И. Н. Канонические представления случайных функций и их применение в задачах компьютерной поддержки научных исследований. — М.: ТОРУС ПРЕСС, 2009. 768 с.
11. Синицын И. Н., Синицын В. И., Корепанов Э. Р., Белоусов В. В., Сергеев И. В., Басилашвили Д. А. Опыт моделирования эредитарных стохастических систем // Кибернетика и высокие технологии XXI века: Сб. докл. XIII Междунар. науч.-технич. конф. — Воронеж: Саквоее, 2012. Т. 2. С. 346–357.
12. Синицын И. Н. Параметрическое статистическое и аналитическое моделирование распределений в нелинейных стохастических системах на многообразиях // Информатика и её применения, 2013. Т. 7. Вып. 2. С. 4–16.
13. Синицын И. Н. Сергеев И. В., Синицын В. И., Корепанов Э. Р., Белоусов В. В. Математическое обеспечение параметрического моделирования распределений в интегродифференциальных стохастических системах // Системы и средства информатики, 2014. Т. 24. № 1. С. 4–45.
14. Градштейн И. С., Рыжик И. М. Таблицы интегралов, сумм, рядов и произведений. — М.: ГИФМЛ, 1963. 1100 с.

*Поступила в редакцию 13.05.14*

## MATHEMATICAL SOFTWARE TOOLS FOR ANALYTICAL MODELING IN STOCHASTIC SYSTEMS WITH COMPLEX NONLINEARITIES

**I. N. Sinitsyn, V. I. Sinitsyn, I. V. Sergeev, E. R. Korepanov, V. V. Belousov,  
and V. S. Shorgin**

Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str.,  
Moscow 119333, Russian Federation

**Abstract:** New mathematical software tools in “IDStS” are developed for differential stochastic systems (DStS) with Wiener and Poisson noises with complex finite, differential, and integral nonlinearities and hereditary StS (HStS) reducible to DStS. Algorithms of analytical modeling based on the normal approximation method (NAM) and the statistical linearization method (SLM) are given. Numeral and symbolic algorithms for NAM (SLM) are developed in MATLAB for typical nonlinearities in DStS. Test examples are given. Special attention is paid to the problems of analytical modeling in HStS with complex nonlinearities based on NAM and SLM.

**Keywords:** analytical modeling; complex finite differential and integral nonlinearities; complex irrational nonlinearities; differential stochastic system with Wiener and Poisson noises; mathematical software tool “IDStS;” method of normal approximation; method of statistical linearization; hereditary stochastic systems reducible to differential

**DOI:** 10.14357/08696527140301

### Acknowledgments

The work was partly financially supported by the Program of the RAS Department for Nanotechnologies and Information Technologies “Intelligent information technology, system analysis, and automation” (project 1.7).

### References

1. Sinitsyn, I. N., and V. I. Sinitsyn. 2013. *Lektsii po normal'noy i ellipsoidal'noy approksimatsii raspredeleniy v stokhasticheskikh sistemakh* [Lectures on normal and ellipsoidal approximation of distributions in stochastic systems]. Moscow: TORUS PRESS. 488 p.
2. Sinitsyn, I. N., and V. I. Sinitsyn. 2014 (in press). Analiticheskoe modelirovaniye normal'nykh protsessov v stokhasticheskikh sistemakh so slozhnymi nelineynostyami v stokhasticheskikh sistemakh s razryvnymi kharakteristikami [Analytical modeling of normal processes in stochastic systems with complex nonlinearities]. *Informatika i ee Primeneniya — Inform. Appl.* 8(2). doi: 10.14357/19922264140302.

3. Pugachev, V. S., and I. N. Sinitsyn. 1987. *Stochastic differential systems. Analysis and filtering*. Chichester, New York: Jonh Wiley. 549 p.
4. Pugachev, V. S., and I. N. Sinitsyn. 2001. *Stochastic systems. Theory and applications*. Singapore: Worls Scientific. 908 p.
5. Sinitsyn, I. N. 2014. Analiz i modelirovanie raspredeleniy v ereditarnykh stokhasticheskikh sistemakh [Analysis and modeling of distributions in hereditary stochastic systems]. *Informatika i ee Primeneniya — Inform. Appl.* 8(1):2–11.
6. Voronov, A. A. 1979. *Ustoychivost', upravlyayemost', nablyudayemost'* [Stability, controllability, observability]. Moscow: Nauka. 336 p.
7. Voronov, A. A. Vvedenie v dinamiku slozhnykh upravlyayemykh sistem [Introduction to complex control systems dynamics]. Moscow: Nauka, 1985. 352 p.
8. Popov, B. A., and G. S. Tesler. 1984. *Vychislenie funktsiy na EVM: Spravochnik* [Computing of functions]. Kiev: Naukova Dumka. 599 p.
9. Korn, G. A., and T. M. Korn. 1968. *Mathematical handbook*. New-York: McGraw-Hill.
10. Sinitsyn, I. N. 2009. *Kanonicheskie predstavleniya sluchaynykh funktsiy i ikh prime-nenie v zadachakh kompyuternoy podderzhki nauchnykh issledovaniy* [Canonical expansions of random functions and their application to scientific computer-aided support]. Moscow: TORUS PRESS. 768 p.
11. Sinitsyn, I. N., V. I. Sinitsyn, E. R. Korepanov, V. V. Belousov, I. D. Sergeev, and D. A. Basilashvili. 2012. Opyt modelirovaniya ereditarnykh stokhasticheskikh sistem [Experience of modeling in hereditary stochastic systems]. *Kibernetika i Vysokie Tekhnologii XXI Veka: Sb. dokl. XIII Mezhdunar. nauch.-tehnich. konf.* [Cybernetics and High Technologies of XXI Century: 13th Scientific and Technological Conference (International) Proceedings]. Voronezh: Sakvoee. 2:346–357.
12. Sinitsyn, I. N. 2013. Parametricheskoe statisticheskoe i analiticheskoe modelirovanie raspredeleniy v nelineynykh stokhasticheskikh sistemakh na mnogoobraziyakh [Parametric statistical and analytical modeling of distributions in stochastic systems on manifolds]. *Informatika i ee Primeneniya — Inform. Appl.* 7(2):4–16.
13. Sinitsyn, I. N., I. V. Sergeev, V. I. Sinitsyn, E. R. Korepanov, and V. V. Belousov. 2014. Matematicheskoe modelirovanie parametricheskogo model'rovaniya raspredeleniy v integrodifferentsial'nykh stokhasticheskikh sistemakh [Distributions parametrical modeling software for integrodifferential stochastic systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(1):4–45.
14. Gradshteyn, I. S., and I. M. Ryzhik. 1963. *Tablitsy integralov, summ, ryadov i proizvedeniy*. [Tables of integrals, sums, series, and products]. Moscow: GIFML. 1100 p.

Received May 13, 2014

## Contributors

**Sinitsyn Igor N.** (b. 1940)— Doctor of Science in technology, professor, Honored scientist of RF, Head of Department, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sinitzin@dol.ru

**Sinitsyn Vladimir I.** (b. 1968) — Doctor of Science in physics and mathematics, associate professor, Head of Department, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vsinitsin@ipiran.ru

**Sergeev Igor V.** (b. 1965) — Candidate of Science (PhD) in technology, Deputy Director, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; isergeev@ipiran.ru

**Korepanov Eduard R.** (b. 1966) — Candidate of Science (PhD) in technology, Head of Laboratory, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ekorepanov@ipiran.ru

**Belousov Vasiliy V.** (b. 1977) — Candidate of Science (PhD) in technology, Head of Laboratory, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vbelousov@ipiran.ru

**Shorgin Vsevolod S.** (b. 1978) — Candidate of Science (PhD) in technology, senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vshorgin@ipiran.ru

## БЕЗОПАСНЫЕ АРХИТЕКТУРЫ РАСПРЕДЕЛЕННЫХ СИСТЕМ

*A. A. Грушо<sup>1</sup>, Н. А. Грушо<sup>2</sup>, Е. Е. Тимонина<sup>3</sup>, С. Я. Шоргин<sup>4</sup>*

**Аннотация:** Распределенные компоненты корпоративной информационной системы (КИС) взаимодействуют между собой через сеть, обеспечивая информационное сопровождение производственной деятельности предприятия или организации. Существует целый класс уязвимостей, создающих угрозы правильному и безопасному выполнению КИС своих функций. Часто злоумышленники ищут уязвимости, используя вредоносный код, который, попадая в наименее защищенные узлы, автономно распространяется между хостами информационной системы (ИС). Пусть  $V$  — множество компонентов распределенной ИС,  $E$  — множество возможных взаимодействий между компонентами, тогда  $G = (V, E)$  — архитектура распределенной ИС. Рассматривается следующая декомпозиция архитектуры распределенной КИС на иерархию архитектур классов компонентов. Архитектура сети  $A_1$  распределенной ИС представляет собой нижний уровень декомпозиции. Следующие уровни иерархии — это архитектура задач  $A_2$  и информационные технологии (ИТ)  $A_3$ . Верхним уровнем иерархии является архитектура распределенной ИС. Архитектурные угрозы — это транзит через хосты вредоносного кода и утечка информации через ИТ. Безопасной архитектурой следует считать такую архитектуру, которая в рамках доступных средств снижает риски реализации архитектурных угроз. Архитектура ИС может считаться безопасной, если в ней нет прямых взаимодействий между высокорисковыми и высокоцennыми компонентами. Найдены необходимые и достаточные условия существования безопасных архитектур в распределенных КИС. Установлено, при каких условиях в распределенную ИС можно безопасно вносить новые задачи или ИТ или другие изменения.

**Ключевые слова:** информационная безопасность распределенной системы; виртуальные системы; риски; ценные информационные ресурсы; архитектура распределенной информационной системы

**DOI:** 10.14357/08696527140302

### 1 Введение

Распределенные КИС содержат большое количество основного и вспомогательного оборудования, базового и прикладного программного обеспечения (ПО), информационных ресурсов (ИР). Распределенные компоненты КИС взаимодей-

<sup>1</sup>Институт проблем информатики Российской академии наук, факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова, grusho@yandex.ru

<sup>2</sup>Институт проблем информатики Российской академии наук, info@itake.ru

<sup>3</sup>Институт проблем информатики Российской академии наук, eltimon@yandex.ru

<sup>4</sup>Институт проблем информатики Российской академии наук, sshorgin@ipiran.ru

ствуют между собой через сеть, обеспечивая информационное сопровождение производственной деятельности предприятия или организации.

С неоднородностями ИР и программно-аппаратных систем КИС, наличием у некоторых компонентов внешних взаимодействий связан целый класс уязвимостей, создающих угрозы правильному и безопасному выполнению КИС своих функций. Причинами уязвимостей могут быть как естественные ошибки и трудности организации безопасной работы сложной системы, так и деятельность злоумышленников, целенаправленно ищущих уязвимые места КИС с целью нанесения ущерба. Часто злоумышленники ищут уязвимости, используя вредоносный код, который, попадая в наименее защищенные узлы, автономно распространяется между хостами КИС.

Однако использование вирусов, червей и другого вредоносного кода является сложной задачей. Чаще всего для распространения вредоносного кода используются ошибки в конфигурации сети. Доступ вредоносного кода к ценным ИР также предполагает решение ряда сложных задач. Сюда входит обход средств защиты, повышение привилегий, блокирование антивирусов, активизация предметно-ориентированных модулей для поиска ценной информации. Получение команды на уничтожение ценного ИР или скрытая передача ценной информации во внешнюю (глобальную) сеть предполагают сложную организацию скрытого сетевого взаимодействия.

Почти все эти шаги можно предотвратить правильной организацией взаимодействий компонентов КИС. В работе безопасная организация взаимодействий между компонентами КИС называется безопасной архитектурой ИС. Для построения безопасной архитектуры делается следующее предположение. Любые взаимодействия в КИС происходят либо в хосте, либо через сеть. Взаимодействие таких сложных компонентов КИС, как задачи и ИТ, представляет собой взаимодействие множества хостов через сеть и взаимодействие с ИР внутри хостов. С точки зрения безопасной архитектуры КИС можно предположить, что взаимодействие внутри хостов организовано безопасно. Поэтому основная задача, которая решается в работе, — это безопасное взаимодействие между компонентами КИС через сеть.

В теоретических исследованиях проблем информационной безопасности разработаны примеры гарантированной защиты КИС от упомянутых угроз [1–3]. Однако высокая цена гарантированной защиты в большинстве случаев не оправдывается ценностями, обрабатываемыми в КИС. В данной работе предложены методы защиты, которые можно реализовать значительно более дешевыми средствами. В некоторых случаях можно рассчитывать на массовое внедрение предложенных методов построения безопасных архитектур.

## 2 Модели архитектур распределенных систем

Распределенные ИС опираются на вычислительные сети, представляющие собой хосты и каналы между ними. На хостах установлены приложения. Рас-

пределенная система реализует информационные взаимодействия приложений, расположенных на хостах, использующих сеть. Сетевое оборудование, хосты, приложения, ИР, ИТ и др. будем называть компонентами распределенных ИС.

Компьютерная система, использующая виртуализацию, представляет собой физический сервер (клUSTER), на котором установлено семейство виртуальных серверов и других компонентов, взаимодействующих между собой. С логической точки зрения распределенная ИС и виртуальная ИС идентичны, и далее будем их отождествлять.

Определим понятие архитектуры. Пусть  $V$  — множество компонентов распределенной ИС,  $E$  — множество возможных взаимодействий между компонентами, тогда  $G = (V, E)$  — архитектура распределенной ИС. Можно рассматривать архитектуру распределенной ИС, зависящую от времени:  $G_t = (V_t, E_t)$ , что соответствует ИТ. Для простоты далее будем рассматривать архитектуру распределенной ИС, не зависящую от времени.

Если взаимодействия рассматриваются только попарно и в аспекте «есть или нет», то архитектура представляет собой ориентированный граф, в котором дуга идентифицирует возможность существования некоторого информационного потока от одного компонента к другому. В общем случае взаимодействие — это сложная структура, которая сама может содержать свои компоненты и взаимосвязи между ними. Аналогично компоненты могут иметь сложную структуру, также состоящую из компонентов и взаимосвязей между ними.

Для анализа сложных систем, как правило, используют иерархическую декомпозицию. При исследовании архитектур распределенных ИС также имеет смысл воспользоваться иерархической декомпозицией распределенной и/или виртуальной системы. Далее рассматривается следующая декомпозиция архитектуры распределенной ИС на иерархию архитектур классов компонентов.

Архитектура сети  $A_1$  распределенной ИС представляет собой нижний уровень декомпозиции. Компоненты сети образуют множество хостов  $V_1$ , взаимодействие этих компонентов отождествляется с логическими каналами связи  $E_1$  между хостами. На хостах расположены ИР, которые также имеют компонентную структуру, но не могут самостоятельно взаимодействовать. Информационные ресурсы передаются при взаимодействии хостов.

Следующий уровень иерархии — это архитектура задач  $A_2$ . Каждая задача определяется сбором информации, ее обработкой и распределением. Уровень сети предоставляет задаче информацию, собранную из различных хостов, и услугу распределения информации в другие хосты. Хосты предоставляют задаче услугу в виде ИР и услугу обработки информации. Итак, компонентами архитектуры задач являются хосты, содержащие или обрабатывающие ИР, а взаимодействие хостов, участвующих в сборе, обработке и распределении ИР, определяет взаимодействия в архитектуре задач. В компьютерной практике задачу часто отождествляют с приложением.

Следующим уровнем иерархии  $A_3$  является ИТ. Информационную технологию можно представлять как последовательность задач, выполняемых в интересах

единой цели. Компонентами архитектуры ИТ являются задачи, множество взаимодействий в архитектуре ИТ определяется взаимодействием последовательности решаемых задач.

Верхним уровнем иерархии является архитектура распределенной ИС. Архитектуру распределенной ИС можно рассматривать как множество ИТ. Компонентами архитектуры распределенной ИС являются ИТ, множество взаимодействий ИТ определяет множество взаимодействий в распределенной ИС.

Таким образом, задача представляет собой множество хостов и взаимодействий этих хостов, т. е. логических каналов между этими хостами. Информационная технология представляет собой последовательность задач и их взаимодействие между собой. Иными словами, архитектура ИТ — это также множество хостов и логических каналов, являющееся объединением соответствующих множеств архитектур задач. При этом не возникает дополнительных взаимодействий, так как взаимосвязь задач в ИТ соответствует распределению информации и сбору информации следующих друг за другом задач. Из этих рассуждений следует, что  $A_1$  является основной архитектурой, а все следующие над ней в качестве компонентов имеют подмножества  $V_1$  и взаимосвязи подмножества  $E_1$ . Новая задача или ИТ может увеличить  $E_1$ , но общая схема останется той же.

Рассмотрим некоторые примеры архитектур.

## 2.1 Клиент-серверная архитектура

Клиент-серверная архитектура поддерживается сетью с топологией «звезды»: основные приложения находятся на сервере, пользователи с помощью рабочих станций обращаются к серверным приложениям через сеть коммуникаций. В данной модели могут также присутствовать сетевые устройства, которые для простоты здесь опущены.

Развитием клиент-серверной архитектуры являются ERP (Enterprise Resource Planning) системы, в которых присутствует множество серверов и реализован шлюз, через который хосты могут обращаться к серверной системе.

Примером реализации шлюза является веб-портал. Взаимодействие пользователей с серверной группой осуществляется с помощью веб-взаимодействия или браузера.

## 2.2 VLAN-архитектура

Архитектура сети взаимодействий рабочей группы может быть реализована через VLAN (Virtual Local Area Network), которая реализуется через коммутатор, разрешающий взаимодействие всем членам рабочей группы. Выход за пределы такой сети осуществляется с помощью маршрутизатора.

Корпоративная сеть на базе VLAN позволяет реализовать разбиение корпоративной системы на изолированные домены.

### **3 Архитектурные угрозы**

#### **3.1 Транзит вредоносного кода**

Если вредоносный код был внесен на один из хостов корпоративной сети, а целью вредоносного воздействия является ценная информация на другом хосте, то, используя взаимодействие приложений и лежащую под ней сетевую архитектуру, вредоносный код может распространяться до нужного хоста и реализовать вредоносное воздействие.

Примером такой угрозы является известный червь Stuxnet, который, будучи внедренным с помощью флеш-носителя на одну из рабочих станций, достиг системы управления центрифугами обогащения урана в Иране и вывел их из строя.

#### **3.2 Утечка информации через информационные технологии**

В ИТ для решения отдельных задач может использоваться ценная информация. В реализации ИТ могут участвовать несколько приложений и хостов, и ценная информация может попасть на хост, к которому имеет доступ злоумышленник. Получение несанкционированного доступа злоумышленника к ценной информации является утечкой этой информации. Например, если с хоста злоумышленника можно получить ftp-доступ к хосту с ценной информацией, то вместе с разрешенной информацией злоумышленник может скачать и запрещенную ценную информацию.

## **4 Задача контроля взаимодействий**

Среди ряда функциональных задач при организации архитектуры распределенной системы или системы виртуализации присутствует задача противодействия архитектурным угрозам. В частности, возникает необходимость контролировать взаимодействия так, чтобы не допустить или максимально затруднить вредоносные воздействия при взаимодействиях компонентов архитектуры. В простейшем случае компоненты более сложного взаимодействия — это структурные элементы получения, отправки, преобразования и контроля информации, которую отдают и принимают компоненты низкого уровня архитектуры КИС. Взаимодействия сложных компонентов в реализации взаимодействия низкого уровня могут создать или, наоборот, затруднить реализации угроз, возникающих при низком уровне взаимодействий. Безопасной архитектурой следует считать такую архитектуру, которая в рамках доступных средств снижает риски реализации архитектурных угроз.

## **5 Модель контроля взаимодействий**

Формализуем понятие безопасной архитектуры. Будем считать, что на множестве  $V$  компонентов распределенной системы определено понятие опасности

доступа злоумышленника. Например, на компоненте проще закрепиться временноносному коду. Для сокращения будем называть понятие опасности доступа злоумышленника риском. При этом некоторые компоненты обладают большим риском, некоторые меньшим. Например, компоненты, связанные с Интернетом, обладают большим риском, чем компоненты, не связанные с Интернетом. Аналогично понятие рисков вводится и на множестве взаимодействий  $E$ . Например, канал, проходящий по защищенной территории, имеет меньший риск, чем канал, проходящий по незащищенной территории. Таким образом, имеется некоторая линейная иерархия  $L_1 = \{l_1 < \dots < l_k\}$ , в которую отображаются все компоненты архитектуры. Обозначим это отображение через  $f$  и будем называть его функцией риска,  $f : V_1 \cup V_2 \cup V_3 \cup V_4 \rightarrow L_1$ . Функция  $f$  определена на хостах и однозначно продолжается на множества  $V_2$ ,  $V_3$  и  $V_4$ . В решении каждой задачи могут использоваться ресурсы различных хостов. Риск задачи определяется как максимальное значение риска хостов, взаимодействующих в решении задачи.

Аналогично технология приобретает значение риска, исходя из максимума рисков задействованных в ней задач.

Кроме того, все компоненты категорируются по своей ценности, т. е. имеется некоторая линейная иерархия  $L_2 = \{m_1 < m_2 < \dots < m_s\}$ , в которую отображаются все компоненты архитектуры, включая компоненты ИР. Обозначим это отображение через  $g$  и будем называть его функцией ценностей,  $g : V_{\text{ИР}} \cup \dots \cup V_4 \rightarrow L_2$ . Например, компоненты, в которых находятся криптографические ключи или ключи электронной подписи, являются более ценными, чем компоненты с открытой справочной информацией. Необходимо отметить, что функция  $g$  ценности информации должна быть определена для всех ИР. Уровень ценности компонента естественно определяется как максимум ценности ИР, находящихся или используемых в этом компоненте. Таким образом, все компоненты иерархии приобретают ценность. Уровень ценности нового ИР определяется, как правило, уровнем ценности в том минимальном в иерархической декомпозиции компоненте, в котором этот ресурс создан.

На основании вышеизложенного *концепция безопасности архитектуры* состоит в минимизации или устраниении взаимодействий высокорисковых компонентов с высокоценными компонентами.

Пусть  $f_0$  и  $g_0$  — некоторые пороговые значения из  $L_1$  и  $L_2$  соответственно. Тогда в безопасной архитектуре ИС запрещены взаимодействия компонентов, имеющих  $f \geq f_0$  и  $g \geq g_0$ . Взаимодействия, не являющиеся запрещенными, будем называть *безопасными взаимодействиями*.

С помощью пороговых уровней определяются высокорисковые и высокоценные компоненты распределенной системы. Архитектура ИС может считаться безопасной, если в ней нет прямых взаимодействий между высокорисковыми и высокоценными компонентами, т. е. высокие значения  $f$  и высокие значения  $g$  для взаимодействующих компонентов недопустимы.

Однако иногда возникает необходимость такого неразрешенного взаимодействия. Например, хост обладает ИР высокой и низкой ценности. Несмотря на то

что хост при этом может иметь высокий уровень ценности, ИР в нем, имеющие низкую ценность, могут быть переданы другим хостам с высоким риском. Если такое взаимодействие осуществляется прямо между хостами, то это нарушает основной принцип информационной безопасности архитектуры. Однако можно ввести промежуточное звено (сервер безопасности), которое может оценить уровень передаваемой информации, и если этот уровень достаточно низкий, то сервер безопасности может передать этот ресурс хосту с высоким риском. Если работу сервера безопасности можно считать правильной, обладающей полнотой и непротиворечивостью, то утечки ценной информации к высокорисковым хостам быть не может.

Сервер безопасности может выступать также как средство снижения риска. При этом передача ИР от высокорискового хоста в хост высокой ценности через сервер безопасности может быть признана допустимой.

## **6 Модели взаимодействия высокооцененных компонентов с высокорисковыми компонентами**

Как отмечалось выше, некоторые запрещенные взаимодействия можно осуществлять через серверы безопасности. Рассмотрим примеры таких взаимодействий при использовании сервера безопасности.

### **6.1 Гомоморфизмы ценной информации**

Ценность информации может снижаться, если рассматривать некоторые функции от ценных ИР. Широко известно сопровождение текстового ИР двумя функциями:

- (1) заголовок ценного ИР может иметь меньшую ценность, чем весь текст;
- (2) аннотация ценного текстового ресурса может иметь меньшую ценность, чем исходный ресурс.

Ценность ИР может быть представлена в виде метки, неотделимой от ресурса. Аналогично функции от ИР (заголовок, аннотация) могут иметь отдельные метки ценности. Впервые использование меток, неотделимых от ресурса, было предложено в [1]. На хосте может находиться непомеченная информация, которая при передаче приобретает ценность хоста.

При обращении высокорискового компонента к высокооцененному компоненту за высокоценным ИР в ответ может быть предложена функция от исходного ресурса, имеющая меньшую ценность. Сервер безопасности проверяет допустимость передачи функции от ресурса высокорисковому компоненту. Если такое взаимодействие допустимо по правилам безопасности архитектуры, то эта информация передается высокорисковому компоненту. В противном случае взаимодействие считается недопустимым.

## 6.2 Статистические данные

Широко известным методом построения функции с меньшим значением ценности являются статистические данные от ценных ресурсов. Этот метод часто применяется в медицинских учреждениях для ознакомления широких слоев практикующих врачей с данными по применению новых лекарственных препаратов. Диагноз конкретного человека и особенность протекания его болезни являются ценной информацией. Статистические данные по применению лечебных препаратов не содержат персональных данных и могут передаваться высокорисковым компонентам распределенных систем.

Впервые использование статистики для снижения ценности ИР было предложено в [4].

## 6.3 Шифрование

При использовании высокорисковых каналов связи (также компонентов сети), например проходящих через незащищенную территорию, в качестве метода снижения риска используется шифрование. Ценная информация в зашифрованном виде может передаваться по незащищенному каналу. При этом на приемном конце исходная ценная информация может быть восстановлена без потерь.

В перспективе для вычисления функций можно использовать искусственный интеллект. Например, выделять в ИР части, которые можно использовать при взаимодействии с высокорисковыми компонентами, не нарушая правила безопасной архитектуры.

## 6.4 Снижение рисков

Для безопасного взаимодействия высокорисковых компонентов с высокочастотными можно использовать методы снижения рисков, которые реализуются с помощью сервера безопасности. Приведем пример.

Если высокорисковый компонент обращается к высокочастотному компоненту с запросом на решение ценной задачи, результат решения которой не является ценным, то сервер безопасности может реализовать изолированную среду, куда будут внесены условия с высокорисковых компонентов и ценная информация для решения ценной задачи. Результат может быть передан в высокорисковый компонент, а изолированная среда вместе со всей информацией уничтожена. В этом случае взаимодействие ценных компонентов с высокорисковыми компонентами исключается.

# 7 Снижение рисков для серверов безопасности

Естественно, возникает вопрос о безопасности взаимодействий с сервером безопасности. Один из приемов для существенного затруднения вредоносного воздействия на сервер безопасности в виртуальных средах — это их размещение

на более низких уровнях иерархии виртуальных систем. Например, на уровне гипервизора или нулевого домена. В распределенных системах необходимо изолированное размещение серверов безопасности на отдельных хостах и использование безопасных протоколов взаимодействия. Возможны другие решения.

## 8 Свойства безопасных архитектур

Из множества компонентов собираются новые компоненты не менее высокого уровня в иерархической декомпозиции. Таким образом, любой компонент распределенной ИС с точки зрения безопасной архитектуры состоит из множества хостов, логических каналов между ними и множества ИР. Поэтому функции  $f$  и  $g$  не могут определяться на всех компонентах произвольно, значения каждой из этих функций зависят следующим образом. Пусть компонент  $K$  содержит множество хостов  $\tilde{V}$  и множество взаимодействий хостов  $\tilde{E}$ , тогда

$$f(K) = \max_{x \in \tilde{V}} f(x). \quad (1)$$

Аналогично

$$g(K) = \max_{x \in \tilde{V}} g(x). \quad (2)$$

Компонент  $K$  корректно определен, если все взаимодействия в нем безопасны.

Таким образом, в безопасной архитектуре выполняются следующие правила:

- риск (ценность) корректно определенного компонента равен максимально-му из рисков (ценностей) составных корректно определенных компонентов. Обозначим это правило (\*);
- если для запрещенного взаимодействия двух корректно определенных компонентов риск нельзя снизить или нельзя использовать для взаимодействия ИС более низкой ценности в высокоценном компоненте, то взаимодействие таких компонентов запрещено. Обозначим это правило (\*\*).

Теперь можно написать формальное определение безопасной архитектуры ИС.

**Определение.** Распределенная информационная система имеет безопасную архитектуру тогда и только тогда, когда для любых двух корректно определенных компонентов  $V_i, V_j \in V$  и заданных пороговых значений  $f_0, g_0$ , если  $f(V_i) \geq f_0$ ,  $g(V_j) \geq g_0$ , то выполняется условие (\*\*) и некорректных компонентов в ИС нет.

**Утверждение 1.** Если распределенная ИС строится на базе сети, для которой выполнено условие (\*\*), а значения функций  $f$  и  $g$  для компонентов вычисляются по правилу (\*), то любые компоненты в распределенной ИС корректно определены и для любой пары из них выполняется условие (\*\*).

**Доказательство.** Произвольный компонент  $K$ , содержащий множество хостов  $\tilde{V}$  и множество взаимодействий хостов  $\tilde{E}$ , содержит два компонента  $V_i$

и  $V_j$ . Для каждого из них выполняются условия (1) и (2). Если для максимумов составных компонентов нет запрещенного взаимодействия (по условию (\*\*)), то их нет для хостов, входящих в эти компоненты, и для задействованных ИР. Если компонент  $K$  некорректно определен, то тогда в нем между какими-то компонентами  $V_i$  и  $V_j$  есть запрещенное взаимодействие. Тогда из конечности состава компонентов и взаимодействий должна существовать пара хостов, имеющих запрещенное взаимодействие. Это противоречит условию.

Это рассуждение можно повторить для любой пары компонентов из  $V$ . Поэтому условие (\*\*) в распределенной ИС выполняется.

**Утверждение 2.** *Распределенная ИС обладает безопасной архитектурой тогда и только тогда, когда она строится на базе сети, для которой выполнено условие (\*\*), а значения функций  $f$  и  $g$  для компонентов вычисляются по правилу (\*).*

**Доказательство.** Достаточность следует из утверждения 1 и определения безопасной архитектуры.

*Необходимость.* Выполнение условия (\*) следует из определения безопасной архитектуры. Выполнение условия (\*\*) для сети следует из того, что любые взаимодействия в системе, в том числе в сети, безопасны.

**Следствие 1.** Пусть распределенная ИС обладает безопасной архитектурой. Тогда исключение или добавление новых компонентов оставляет архитектуру системы безопасной.

Следствие показывает, при каких условиях в распределенную ИС можно безопасно вносить новые задачи и ИТ или вносить другие изменения.

## 9 Примеры безопасных архитектур

### 9.1 Простейшая безопасная архитектура на базе виртуализации

В настоящее время многие центральные процессоры поддерживают виртуализацию. Изоляция доменов предполагает изоляцию доступных разделов памяти, изоляцию процессов и потоков, изоляцию взаимодействий с различными устройствами ввода–вывода. В связи с этим возможно создание изолированных виртуальных доменов, разделяющих подготовку платежных документов и работу с другой ценной информацией в одних доменах, а взаимодействие с Интернетом — в других доменах. Поскольку информационный обмен через Интернет является уязвимым (вирусы, вредоносный код и др.), а информация, связанная с установкой электронной подписи, составлением платежных документов, может нанести пользователю большой ущерб, то целесообразно работать с более ценной информацией в одном домене, а с менее ценной информацией — в другом. В том числе при выходе в глобальную сеть целесообразно работать в другом изолированном виртуальном домене, существенно затрудняя доступ к более ценной информации. Такое разграничение доступа необходимо в интересах пользователя.

Так как подобная изоляция происходит на уровне гипервизоров, то в таких приложениях мобильных систем, как дистанционное банковское обслуживание, оплата услуг жилищно-коммунального хозяйства, BYOD, возможно обеспечение достаточной защиты ценной информации от атак из глобальной сети или из несанкционированных носителей памяти.

Вместе с тем данное решение является программным, что определяет его низкую стоимость при массовом применении. Отметим, что каждая отдельная виртуальная операционная система с приложениями небезопасна для атак из сети и съемных носителей. Однако в указанной архитектуре удалось получить общее повышение безопасности совокупной системы, т. е. с помощью архитектуры виртуальных систем можно построить более безопасную систему, чем исходная система.

Установим соответствие с ранее введенными понятиями. Иерархии  $L_1 = L_2$  состоят из двух уровней. Высокий риск — это связь хоста с Интернетом, низкий — отсутствие связи с Интернетом. Высокая ценность ИР — персональные данные, включая электронную подпись, личный архив и др., низкая ценность — общедоступная информация или информация, не представляющая ценности для пользователя. В высокоценном компоненте ценные ИР могут помечаться хешем, который также хранится на сервере безопасности. Сервер безопасности — виртуальная машина, контролирующая появление персональной информации при передаче данных в высокорисковую виртуальную машину для отправки в Интернет. Сервер безопасности также проводит антивирусную проверку в информации, которую пользователь хочет передать в высокоценный компонент.

## **9.2 Системы с многоуровневой политикой безопасности**

Многоуровневая политика безопасности (multilevel security — MLS) основана на решетке ценностей (в частном случае, иерархии) [5]. В системе информационный поток считается допустимым, если он направлен по решетке снизу вверх. MLS является основной политикой защиты конфиденциальности, введена в TCSEC (Trusted Computer System Evaluation Criteria) [1], реализуется с помощью сравнения меток субъекта, запросившего доступ к объекту, и объекта. В реализациях MLS роль сервера безопасности играет монитор обращений, который сравнивает метки субъекта и объекта и по результату сравнения разрешает доступ или нет. MLS является частным случаем описания безопасной архитектуры. Однако MLS требует защиты от скрытых каналов, что является сложной задачей в ИС.

## **10 Заключение**

Интенсивное развитие инфраструктуры информатики требует решения серьезных задач в области обеспечения информационной безопасности. Это в свою очередь заставляет ученых обращаться к теоретическим вопросам защиты информации. Несмотря на то что было построено большое количество формальных

моделей информационной безопасности, по-прежнему существует потребность в разработке простых и понятных моделей информационной безопасности, связанных с быстроразвивающимися распределенными и виртуальными системами.

В статье сделана попытка создания такой модели. Несмотря на сложность современных распределенных ИС, получены достаточно простые условия защищенности таких систем. При этом вариации требований к защите сервера безопасности и его функционал обеспечивают выполнение различных по строгости требований к защите информации.

Дальнейшие исследования в этой области связаны с практическими реализациями изложенного в статье подхода.

## Литература

1. DoD 5200.28-STD: Department of Defense Trusted Computer System Evaluation Criteria. — Department of Defense Standard, December 1985.
2. ГОСТ Р ИСО / МЭК 15408-1-2002. Информационные технологии. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 1. Введение и общая модель. — М.: Изд-во стандартов, 2002.
3. ГОСТ Р ИСО / МЭК 15408-2-2008. Информационные технологии. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 2. Функциональные требования безопасности. — М.: Изд-во стандартов, 2008.
4. Denning D. E., Schlorer J. Inference controls for statistical databases // Computer, 1983. Vol. 1. No. 7. P. 69–82.
5. Грушо А., Применко Э., Тимонина Е. Теоретические основы компьютерной безопасности. — М.: Академия, 2009. 272 с.

*Поступила в редакцию 01.08.14*

---

---

## SECURE ARCHITECTURE OF DISTRIBUTED SYSTEMS

**A. A. Grusho<sup>1,2</sup>, N. A. Grusho<sup>1</sup>, E. E. Timonina<sup>1</sup>, and S. Ya. Shargin<sup>1</sup>**

<sup>1</sup>Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

<sup>2</sup>Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation

**Abstract:** Distributed components of an enterprise information system (IS) interact among themselves through a network, providing information about production activity of an enterprise or an organization. There is a class of vulnerabilities creating threats to correct and safe execution of the functions of IS. Often, hackers look for vulnerabilities using a malicious code which independently extends between hosts of IS while getting to the least protected

nodes. Let  $V$  be the set of components of the distributed IS and  $E$  be the set of possible interactions between components, then  $G = (V, E)$  is the architecture of the distributed IS. The following decomposition of architecture of the distributed IS is considered as a hierarchy of architecture of classes of components. The architecture of the network of the distributed IS represents the bottom level of decomposition. The following levels of hierarchy are the architecture of applications and the architecture of information technologies. The top level of hierarchy is the architecture of the distributed IS. Architectural threats are the transit through hosts of the malicious code and information leakage through information technologies. Secure architecture is such architecture which within available means reduces risks of implementation of architectural threats. The architecture of IS is secured if there are no direct interactions between high-risky and high-valuable components. Necessary and sufficient conditions of existence of secure architecture of a distributed IS are found. It is defined under what condition it is possible to insert new tasks or information technologies to a distributed IS or to make other changes safely.

**Keywords:** information security of a distributed system; virtual system; risks; valuable information resources; architecture of a distributed information system

**DOI:** 10.14357/08696527140302

## References

1. Department of Defense Standard. December 1985. DoD 5200.28-STD Department of Defense Trusted Computer System Evaluation Criteria.
2. ISO / IEC 15408-1-2002. Informatsionnye tekhnologii. Metody i sredstva obespecheniya bezopasnosti. Kriterii otsenki bezopasnosti informatsionnykh tekhnologiy. Chast' 1. Vvedenie i obshchaya model' [Information technology. Security techniques. Evaluation criteria for IT security. Part 1. Introduction and general model]. 2002. Moscow: IPK Standards Publishing House.
3. ISO / IEC 15408-2-2008. Informatsionnye tekhnologii. Metody i sredstva obespecheniya bezopasnosti. Kriterii otsenki bezopasnosti informatsionnykh tekhnologiy. Chast' 2. Funktsional'nye trebovaniya bezopasnosti [Information technology. Security techniques. Evaluation criteria for IT security. Part 2. Security functional requirements]. 2008. Moscow: IPK Standards Publishing House.
4. Denning, D. E., and J. Schlorer. 1983. Inference controls for statistical databases. *Computer* 1(7):69–82.
5. Grusho, A., E. Primenko, and E. Timonina. 2009. *Teoreticheskie osnovy komp'yuternoy bezopasnosti* [Theoretical bases of computer security]. Moscow: Publishing Center Academy. 272 p.

Received August 1, 2014

## Contributors

**Grusho Alexander A.** (b. 1946) — Doctor of Science in physics and mathematics, Corresponding member of the Russian Academy of Cryptography; leading scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov

Str., Moscow 119333, Russian Federation; professor, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; grusho@yandex.ru

**Grusho Nick A.** (b. 1982) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; info@itake.ru

**Timonina Elena E.** (b. 1952) — Doctor of Science in technology, professor, leading scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; eltimon@yandex.ru

**Shorgin Sergey Ya.** (b. 1952) — Doctor of Science in physics and mathematics, professor, Deputy Director, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sshorgin@ipiran.ru

## СКРЫТЫЕ КАНАЛЫ В БЕСПРОВОДНЫХ СЕТЯХ СТАНДАРТА 802.11\*

*H. A. Грушо<sup>1</sup>*

**Аннотация:** Изучается возможность построения скрытых каналов с помощью служебных заголовков сетевых кадров стандарта 802.11, а также некоторых особенностей функционирования беспроводных сетей на базе этого стандарта. Для построения скрытых каналов приведено описание принципов работы беспроводной сети, рассмотрен процесс подключения и отключения клиентов беспроводной сети, указана структура сетевого кадра. Для подтверждения возможности использования заголовков сетевого кадра и механизма деаутентификации клиента беспроводной сети для скрытой передачи данных приведено описание макета, в состав которого входит необходимое программное и аппаратное обеспечение. Результаты экспериментов подтверждают возможность построения скрытых каналов в беспроводных сетях стандарта 802.11. Автором предложены способы использования описанных в статье скрытых каналов, а также методы противодействия им.

**Ключевые слова:** информатика; компьютерные сети; беспроводные сети; информационная безопасность; скрытые каналы

**DOI:** 10.14357/08696527140303

### 1 Введение

Теоретические возможности построения скрытых каналов через стандартные сети между программно-аппаратными агентами в компьютерных системах исследовались во многих работах [1–5]. Данная работа является продолжением практических исследований [6, 7] по построению скрытых каналов с использованием технических особенностей реализации протоколов связи.

В статье рассматривается беспроводная сеть Wi-Fi стандарта IEEE 802.11(a, b, g, n, ac) [8]. Wi-Fi сети организуются с помощью специального оборудования: точки доступа (ТД) и беспроводных адаптеров. Точка доступа производит объединение беспроводных клиентов в единую локальную беспроводную сеть. На клиентах (персональных компьютерах (ПК), планшетах, смартфонах) должен быть установлен беспроводной адаптер — специальный приемопередатчик, позволяющий клиенту работать с беспроводной сетью.

---

\*Работа частично поддержана РФФИ (проект 13-01-00215).

<sup>1</sup>Институт проблем информатики Российской академии наук, info@itake.ru

Для подключения клиента к беспроводной сети необходимо знать название (SSID — service set identifier) беспроводной сети, а также, если включено шифрование беспроводной сети, пароль доступа к ней. Шифрование нужно для того, чтобы защитить трафик беспроводной сети от перехвата злоумышленником. Однако шифруются не все данные, передаваемые по беспроводной сети. Например, служебные заголовки пакета не шифруются и передаются в открытом виде, т. е. любой клиент, не имеющий пароля к беспроводной сети, может прочитать их.

Так как скрытые каналы — это вопрос информационной безопасности, то этим вопросом занимаются многие научные организации и коллектизы. Из открытых источников [9, 10] известно, например, о возможности построения скрытых каналов с помощью переключения скорости передачи данных между клиентом и ТД.

Возможности и способы использования незашифрованных полей для передачи данных в беспроводной сети рассматриваются в этой статье.

В разд. 2 рассматривается стандарт 802.11 и определяются необходимые параметры и инструменты для построения скрытых каналов. В разд. 3 приводится описание макета для проведения экспериментов, а также последовательность проведения эксперимента. Разделы 4–6 описывают проведение экспериментов и их результаты. Возможности применения скрытой передачи данных и методы ее предотвращения описаны в разд. 7.

## 2 Модель

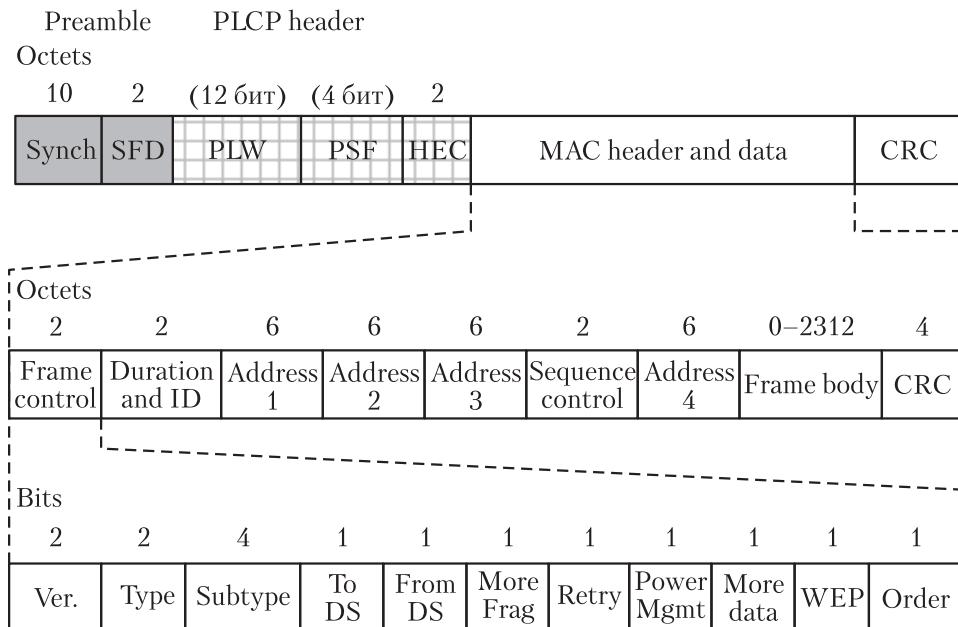
Основная цель исследования — возможность построения скрытых каналов в беспроводных сетях Wi-Fi. Для этого необходимо решить ряд промежуточных задач: изучить описание стандарта 802.11 с точки зрения структуры данных, передаваемых по беспроводной сети, возможных модификаций этих данных и последствий выявления такого рода модификаций, рассмотреть программное обеспечение (ПО), которое позволит анализировать данные, передаваемые по беспроводной сети. Изучив возможности для построения скрытых каналов в Wi-Fi, можно оценить потенциальную пропускную способность таких каналов, их надежность (уничтожаемость), выявляемость [11].

Существует два типа подключений между беспроводными устройствами Wi-Fi: соединение AdHoc и Infrastructure [8]. В случае AdHoc происходит соединение клиент–клиент без участия ТД. В случае Infrastructure клиенты подключаются к ТД и обмениваются данными через нее. В данной статье рассматривается подключение типа Infrastructure, потому что AdHoc-соединения используются крайне редко для решения каких-либо локальных задач, не связанных (обычно) с доступом к глобальным сетям.

Клиент и ТД обмениваются кадрами, которые имеют структуру, описанную [8] в стандарте 802.11. Структура этого пакета приведена на рис. 1.

Полная расшифровка всех полей и их значений — это очень большой документ, поэтому в данной статье будут использованы и расшифрованы лишь

## 802.11 Frame with frequency hopping PHY PLCP



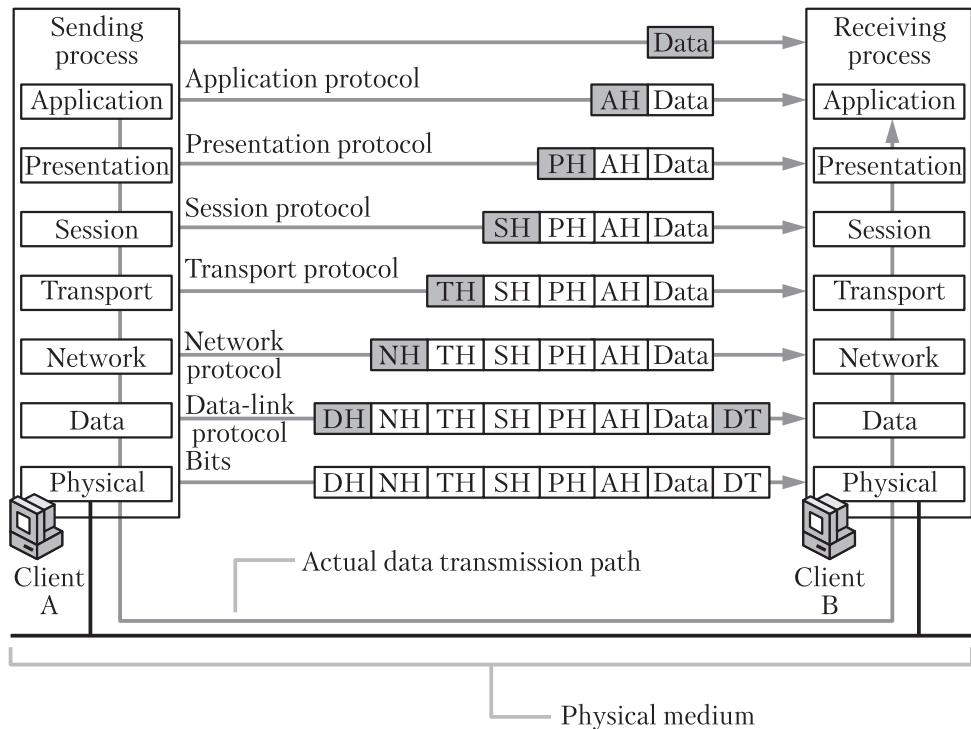
**Рис. 1** Структура кадра 802.11 [12]

те поля, которые необходимы для поставленной задачи. В заголовках сетевых кадров обычно указываются длины данных, контрольные суммы, адреса отправителей, получателей, управляющие команды, номера пактов и т. д.

Беспроводные сетевые адAPTERы, установленные на клиентах, через специальный драйвер взаимодействуют с операционной системой (ОС), а ОС взаимодействует с прикладными программами, предоставляя им возможность сетевого взаимодействия. Беспроводной адAPTER получает сетевой кадр, обрабатывает его и передает инкапсулированные данные в ОС. Модель OSI (Open Systems Interconnection) (рис. 2) описывает принцип инкапсуляции данных на разных уровнях.

Для анализа полученных кадров и отправки новых кадров необходимо специальное ПО. Прикладные программы, использующие программный интерфейс, предоставляемый ОС Windows, не способны получать доступ к данным сетевого кадра, а получают лишь данные более высоких уровней модели OSI. Поэтому специальное ПО должно обращаться к беспроводному адаптеру через драйвер, позволяющий осуществлять прием и передачу сетевых кадров.

Для проведения исследований необходимо ПО, которое может решать следующие задачи:

**Рис. 2** Модель OSI [13]

- взаимодействие с беспроводным адаптером «напрямую»;
- прослушивание беспроводной сети с целью захвата сетевых кадров от клиентов и ТД;
- отображение полученных сетевых кадров;
- формирование новых сетевых кадров;
- передача в беспроводную сеть сформированных или захваченных сетевых кадров.

Исходя из этих условий была выбрана программа CommView [14]. С помощью ПО CommView (ПО CV) можно проводить перехват кадров, передаваемых между клиентом и ТД, анализировать все данные этих кадров, создавать новые кадры необходимой структуры, передавать их в беспроводную сеть. Одной из важных особенностей является тот факт, что для получения и передачи сетевых кадров не требуется знать пароль для беспроводной сети. Отсутствие пароля не позволит прочитать данные более высокого уровня модели OSI, но, как отмечено ранее, не помешает прочитать все служебные заголовки кадра 802.11. Так же и с передачей кадра обратно в беспроводную сеть: невозможно сформировать

данные для более высокого уровня модели OSI, не зная пароль на шифрование данных, но можно изменять необходимым образом служебные заголовки кадра 802.11.

Необходимо обратить внимание на еще одну особенность стандарта 802.11 — процесс подключения клиента к ТД и отключения от нее [8]. Процесс подключения упрощенно происходит следующим образом: клиент обращается к ТД с пробным запросом, ТД отвечает клиенту пробным ответом. Затем клиент производит запрос на аутентификацию, а ТД отвечает на этот запрос. После этого клиент посыпает запрос на ассоциацию. Точка доступа подтверждает ассоциацию. После этой процедуры подключения становится возможной передача зашифрованных данных между клиентом и ТД. В стандарте 802.11 предусмотрен механизм деаутентификации. Точка доступа посыпает специальное оповещение клиенту с указанием причины деаутентификации. Так как аутентификация является необходимым условием для ассоциации клиента, то деаутентификация приводит к завершению ассоциации клиента и ТД. Таким образом, передача данных между клиентом и ТД становится невозможной до проведения повторной аутентификации и ассоциации.

Возможность применения умышленного прерывания связи между клиентом и ТД для построения скрытых каналов необходимо проверить с помощью ПО СV, которое способно не только проводить анализ полученных данных, но и передавать кадры деаутентификации в беспроводную сеть.

С помощью ПО СV необходимо произвести следующие исследования:

- проверить возможность захвата сетевых кадров, передаваемых между клиентом и ТД;
- проверить возможность подсчета сетевых кадров, переданных от клиента;
- проверить возможность приема–передачи сетевого кадра с модифицированными заголовками 802.11;
- проверить возможность использования прерывания связи между клиентом и ТД для скрытой передачи данных.

На основании результатов проведенных экспериментов можно будет сделать заключение:

- о возможности применения заголовков 802.11 для организации скрытых каналов;
- о возможности применения кадров деаутентификации для построения скрытых каналов;
- о пропускной способности и надежности описанных скрытых каналов;
- о возможных других способах использования стандарта 802.11 для построения скрытых каналов.

Для проведения экспериментов необходимо не только специализированное ПО СV, но и оборудование, которое будет выступать в качестве ТД и клиента. Все

вместе (ПО СВ, оборудование клиента и ТД) является макетом для проведения экспериментов. Более подробное описание макета и его функционирование приведено ниже.

### 3 Состав макета

Для проведения эксперимента необходимы следующие элементы: ТД, клиент, ПК с беспроводным адаптером и установленным ПО СВ. Так как сетевое взаимодействие обычно заключается в том, что одна сторона выступает в качестве клиента, а другая — в качестве сервера, то удобно использовать два клиента беспроводной сети, один из которых будет «сервером», а второй «клиентом».

Таким образом, макет состоит из трех ПК и ТД. Используются ПК под управлением ОС Windows 7 с установленными на них адаптерами беспроводной сети Atheros и специальным ПО СВ.

Назовем ПК в составе макета следующим образом: А — источник сообщений, Б — получатель сообщений, Н — сторонний наблюдатель. Персональные компьютеры А и Б подключены к беспроводной сети test, которую организует ТД. Настроим ТД таким образом, чтобы использовалось шифрование WPA2-PSK. Персональный компьютер Н не подключен к беспроводной сети, т. е. не имеет ключа для подключения к беспроводной сети test. На всех ПК запущена программа CommView.

### 4 Перехват сетевых кадров беспроводной сети

На ПК А необходимо запустить стандартную программу cmd.exe, которая входит в состав ОС Windows. После выполнения этой команды откроется окно, предлагающее ввести команду. Необходимо выполнить команду «ping X.X.X.X → t». Утилита ping — это стандартное ПО, входящее в состав ОС Windows, которое отправляет специальные icmp-пакеты на указанный в параметре запуска узел; X.X.X.X — это IP-адрес узла ПК Б. После выполнения этой команды ПК А начнет отправлять пакеты на ПК Б.

На ПК Н в ПО СВ необходимо включить режим сканирования беспроводных сетей, нажав кнопку «Начать захват». Затем необходимо выбрать закладку «Пакеты» для отображения захваченных пакетов. В этом режиме отображается в виде таблицы следующая информация: номер полученного пакета, протокол, MAC-адрес источника, MAC-адрес получателя, IP-адрес отправителя, IP-адрес получателя, время получения пакета, размер пакета, дополнительная информация.

После начала захвата в режиме отображения пакетов ПО СВ появится запись о том, что с MAC-адреса ПК А на MAC-адрес ПК Б отправлен пакет. После этого необходимо переключиться в закладку «Nodes» (узлы). В этом режиме будет отображен список узлов и ТД, пакеты от которых были получены ПК Н. Затем необходимо на MAC-адресе ПК А, который будет в этом списке, кликнуть

правой кнопкой и выбрать пункт меню «Details». После этого откроется окно, в котором будет отображена статистика полученных пакетов ПК А.

Так как ПК А передает пакеты узлу ПК Б, то в строке «Total» можно будет наблюдать непрерывный прирост пакетов, например 1, 2, 3 и т. д. Если на ПК А запустить еще одну команду ping с такими же параметрами, то на ПК Н в окне «Details» можно будет увидеть прирост переданных пакетов уже не на 1, а на 2, т. е. 1, 3, 5 и т. д.

Проведенный эксперимент подтверждает возможность:

- перехвата пакетов от узла ПК А к ПК Б;
- подсчета пакетов между ПК А и ПК Б.

Если на ПК А запустить, например, одновременно 10 команд ping с одинаковыми параметрами, то ПК Н сможет увидеть прирост отправленных с ПК А на ПК Б пакетов на 10. Таким образом, выполнение стандартной команды ping на ПК А дает возможность влиять на количество перехваченных ПК Н пакетов от ПК А к ПК Б.

Вариация числа одновременно запущенных команд ping дает возможность построения скрытого канала.

## **5 Передача и перехват модифицированных пакетов**

Программное обеспечение СV позволяет создавать произвольные сетевые пакеты и передавать их в беспроводную сеть. Для передачи данных от ПК А к ПК Н, который не подключен к беспроводной сети, необходимо создать сетевой пакет с модифицированным заголовком, передать его в беспроводную сеть, затем перехватить его на ПК Н и извлечь из перехваченного пакета нужные данные.

На ПК А в ПО СV с помощью пункта меню «Packet Builder» (конструктор пакетов) сделаем произвольный пакет, но укажем в нем нужные MAC-адреса: bssid 20:AA:4B:21:80:4A — идентификатор беспроводной сети, dst 21:21:21:21:21:21, src 20:20:20:20:20:20. Будем считать, что MAC-адреса dst и src — это данные, которые ПК А хочет передать ПК Б. Затем этот пакет передается в беспроводную сеть с помощью кнопки «Send». В это время на ПК Н этот пакет перехватывается с целью извлечения полезных данных, т. е. MAC-адресов получателя и отправителя в кадре.

В таблице полученных пакетов на ПК Н можно увидеть указанные на ПК А MAC-адреса отправителя и получателя: dst 21:21:21:21:21:21, src 20:20:20:20:20:20.

Такой способ передачи данных, в отличие от способа, описанного в разд. 4, требует установленного и запущенного ПО СV на ПК А, но при этом позволяет передавать данные с помощью заголовков 802.11. Следовательно, при модификации хотя бы одного байта MAC-адреса, например получателя, скорость передачи данных составит 1 байт / пакет.

Проведенный эксперимент показывает возможность скрытой передачи данных от ПК А к ПК Н.

## 6 Передача данных в беспроводную сеть

Ранее были описаны два способа передачи данных от ПК А к ПК Н. Рассмотрим обратную задачу: передачу данных от ПК Н, не подключенного к беспроводной сети, к ПК А, который работает в беспроводной сети test.

Для организации такой передачи можно воспользоваться методом деаутентификации узла от беспроводной сети. Для этого ПК Н в ПО СV необходимо сформировать специальный сетевой пакет деаутентификации и передать его в беспроводную сеть. Получив этот пакет, узел будет считать, что ему необходимо повторно пройти процедуру аутентификации и ассоциации в беспроводной сети, что приведет к прерыванию связи. Так как продолжительность работы узла между прерываниями связи и частоту прерывания связи можно измерить на ПК А, то это можно использовать как канал передачи данных от ПК Н к ПК А.

Пусть ПК А выполняет команду «ping x.x.x.x - t», где x.x.x.x — это IP-адрес ПК Б, и ведет подсчет числа запросов, на которые он получил ответ от ПК Б. Персональный компьютер Н передает специальный кадр деаутентификации с помощью ПО СV на ПК А, вызывая тем самым временнéе прерывание связи ПК А и ТД, организующей беспроводную сеть test. Персональный компьютер А на время повторной аутентификации и ассоциации с беспроводной сетью не может получать данные, поэтому ПК А не получит ответа на істmr-запрос выполняемой команды ping. В это время необходимо сохранить на ПК А значение счетчика переданных істmr-пакетов, получивших ответ, а затем очистить его для нового подсчета. После возобновления приема откликов на команду ping от ПК Б производится подсчет пакетов. Затем необходимо повторно отправить пакет деаутентификации с ПК Н. После ассоциации ПК А с беспроводной сетью test необходимо сохранить значение счетчика, а затем очистить его для следующего подсчета. Повторяя эту процедуру несколько раз, можно получить последовательность чисел, которая является данными, передаваемыми от ПК Н к ПК А.

Проведенный эксперимент с прерыванием связи и измерением числа пакетов показывает возможность передачи данных от ПК Н к ПК А.

## 7 Применение описанных способов передачи данных

Допустим, что на ПК А существует вредоносное ПО, которое для простоты будем называть агентом. Для реализации передачи данных, описанной в разд. 4, необходимо, чтобы агент умел преобразовывать данные в частоту передачи пакетов в беспроводную сеть. Если рассматривать вариант, описанный в разд. 5, то агенту на ПК А необходимо внедрить свой драйвер, аналогичный драйверу ПО СV, который сможет работать с беспроводным адаптером на ПК А напрямую.

В этом случае существует вероятность того, что адаптер, установленный на ПК А, не будет поддерживаться драйвером агента, и тогда передача модифицированных пакетов станет невозможной. Более того, агенту для внедрения драйвера потребуются права администратора. Отсутствие прав приведет к невозможности такой передачи данных.

Для реализации приема данных, описанного в разд. 6, агенту не требуется права администратора. Агент должен лишь производить наблюдение за прерываниями связи и продолжительностью беспрерывной работы.

Во всех описанных вариантах приема и передачи данных предполагается, что агент на ПК Н обладает всем необходимым: совместимым с оборудованием драйвером беспроводного адаптера, правами администратора.

Оказать противодействие агентам, внедряющим драйверы и требующим прав администратора, могут как встроенные механизмы защиты ОС Windows, так и антивирусное ПО. Выявленная замена драйвера беспроводного адаптера приводит к невозможности реализации агентом одного из способов приема–передачи, а также выявляет самого агента.

Установка и настройка межсетевых экранов также может заблокировать работу агента. Например, запрет на прием и передачу icmp-пакетов полностью закроет агенту возможность передачи данных, описанных в разд. 4 и 6. Обходом межсетевого экрана может послужить использование другого протокола. К примеру, предполагается, что отличить обращение агента по протоколу **http** к какому-либо серверу от обращения пользователя к аналогичному серверу с помощью браузера более трудоемкая задача для межсетевого экрана, требующая создания списка доверенных приложений, имеющих доступ к компьютерной сети вообще.

Необходимо также отметить, что для защиты беспроводной сети от работы описанных агентов можно использовать ПК, аналогичный ПК Н. Профессиональный системный администратор может выявлять модифицированные пакеты с помощью ПО **CV**, настроив соответствующим образом фильтры сетевых кадров. Но отследить пульсации трафика, создаваемые агентом, без ПО, выявляющего статистическими методами отклонения работы сети, скорее всего, не получится. Более того, если предположить, что агент способен применять разные модификации описанных способов приема–передачи в различных комбинациях, то выявление работы агента становится более трудоемкой задачей. Проведение анализа выявляемости, надежности, уничтожаемости описанных способов приема–передачи — это работа, которая будет описана в следующей статье, посвященной этой тематике.

## **8 Заключение**

В статье рассмотрены способы организации скрытого приема–передачи данных на базе стандарта 802.11. Наличие в заголовке сетевых кадров незашифрованных данных позволяет использовать кадры для передачи произвольных

данных. Свободное распространение электромагнитного излучения беспроводной сети делает возможным перехват сетевого трафика узлом, не подключенным к беспроводной сети. Использование этих особенностей беспроводных сетей позволило построить методы приема–передачи, которые были описаны в статье.

Детальное изучение возможностей стандарта 802.11 приводит к тому, что открываются дополнительные возможности по построению скрытых каналов, таких как, например, использование кадров деаутентификации.

## Литература

1. Грушо А. А. Скрытые каналы и безопасность информации в компьютерных системах // Дискретная математика, 1998. Т. 10. Вып. 1. С. 3–9.
2. Грушо А. А. О существовании скрытых каналов // Дискретная математика, 1999. Т. 11. Вып. 1. С. 24–28.
3. Грушо А. А., Тимонина Е. Е. Языки в скрытых каналах // Информационные технологии в науке, образовании, телекоммуникации, бизнесе: Тр. Междунар. конф. (весенняя сессия). Запорожье: Лаборатория издательских технологий и компьютерной графики Запорожского гос. ун-та. С. 181–184.
4. Грушо А. А., Тимонина Е. Е. Преодоление защиты от скрытых каналов // Обозрение прикладной промышленной математики, 2003. Т. 10. Вып. 3. С. 638–639.
5. Grusho A., Kniazev A., Timonina E. Detection of illegal information flow // Computer network security / Eds. V. Gorodetsky, I. V. Kotenko, V. A. Skormin. Lecture notes in computer science ser., 2005. Vol. 3685. P. 235–244.
6. Грушо Н. А. Скрытые каналы, основанные на метках // Системы и средства информатики, 2013. Т. 23. № 1. С. 7–13.
7. Грушо А. А., Грушо Н. А., Тимонина Е. Е. Скрытые каналы, порожденные метками, в дейтаграммах // Системы и средства информатики, 2013. Т. 23. № 2. С. 3–18.
8. IEEE Standard for Information Technology. Telecommunications and information exchange between systems. Local and metropolitan area networks. Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.
9. Frikha L., Trabelsi Z. A new covert channel in WiFi networks // Risks and Security of Internet and Systems (CRISSIS '08): 3rd Conference (International) Proceedings. — IEEE, 2008. P. 255–260. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4757487>.
10. Calhoun, Jr., T. E., Cao X., Li Y., Beyah R. An 802.11 MAC layer covert channel // Wireless communications and mobile computing. — John Wiley & Sons, 2010. <http://www.cs.gsu.edu/yli/papers/WCMC10.pdf>.
11. Грушо А. А., Грушо Н. А., Тимонина Е. Е. Методы защиты информации от атак с помощью скрытых каналов и программно-аппаратных агентов в распределенных системах // Вестник РГГУ. Сер. Информатика. Защита информации. Математика, 2009. Вып. 10. С. 33–45.
12. Wireless LAN // Data Network Resource. <http://www.rhyshaden.com/wireless.htm>.

13. Data flow in the OSI model. — Microsoft, 2014. <http://technet.microsoft.com/en-us/library/cc977591.aspx>.
14. Tamosoft. <http://www.tamos.ru/products/commview/>.

*Поступила в редакцию 01.08.14*

---

---

## COVERT CHANNELS ON THE WIRELESS NETWORKS OF THE STANDARD 802.11

*N. A. Grusho*

Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str.,  
Moscow 119333, Russian Federation

**Abstract:** The paper studies the possibility of building a covert channel using the header of the 802.11 standard frame and some features of functioning of wireless networks based on these standards. To construct a covert channel, the principles of a wireless network, the connection and disconnection processes of wireless clients, and the structure of the network frame have to be described. To confirm the possibility of using the header of the network frame and the mechanism of deauthentication of a wireless client for covert communication, the paper describes the layout, which includes the necessary hardware and software, and experiments. The experimental results confirm the possibility of building covert channels in 802.11 wireless networks. Authors propose a way of using covert channels described in the article as well as the methods of eliminating them.

**Keywords:** computer science; computer networks; wireless networks; information security; covert channels

**DOI:** 10.14357/08696527140303

### Acknowledgments

The paper was partially supported by the Russian Foundation for Basic Research (project 13-01-00215).

### References

1. Grusho, A. 1998. Hidden channels and information security in computer systems. *Discrete Math. Appl.* 8(2):127–133.
2. Grusho, A. 1999. On the existence of hidden channels. *Discrete Math. Appl.* 9(1):1–5.
3. Grusho, A. A., and E. E. Timonina. 2003. Yazyki v skrytykh kanalakh [Languages in the covert channels]. *Tr. Mezhdunar. Konf. “Informatsionnye Tekhnologii v Nauke, Obrazovanii, Telekommunikatsii, Biznese (vesennyaya sessiya)”* [Conference (International)“Information Technologies in Science, Education, Telecommunications, Business (Spring Session)” Transactions]. Zaporozh'e: Laboratory of Publishing Technologies and Computer Graphics, Zaporozh'e State University. 181–184.

4. Grusho, A. A., and E. E. Timonina. 2003. Preodolenie zashchity ot skrytykh kanalov [Overcoming of protection against the covert channels]. *Obozrenie prikladnoy promyslennoy matematiki* [Review of Applied Industrial Mathematics] 10(3):638–639.
5. Grusho, A., A. Kniazev, and E. Timonina. 2005. Detection of illegal information flow. *Computer network security*. Eds. V. Gorodetsky, I. V. Kotenko, and V. A. Skormin. Lecture notes in computer science ser. 3685:235–244.
6. Grusho, N. 2013. Skrytye kanaly, osnovанные на метках [Covert channels generated by tags]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 23(1):7–13.
7. Grusho, A., N. Grusho, and E. Timonina. 2013. Skrytye kanaly, порожденные метками, в дейтаграммах [Covert channels generated by tags in datagrams]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 23(2):3–18.
8. IEEE Standard for Information Technology. Telecommunications and information exchange between systems. Local and metropolitan area networks. Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.
9. Frikha, L., and Z. Trabelsi. 2008. A new covert channel in WiFi networks. *Risk and Security of Internet and Systems (CRiSIS'08): 3rd Conference (International) Proceedings*. IEEE. Available at: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4757487> (accessed August 1, 2014).
10. Calhoun, Jr., T. E., X. Cao, Y. Li, and R. Beyah. 2010. An 802.11 MAC layer covert channel. *Wireless communications and mobile computing*. John Wiley & Sons. Available at: <http://www.cs.gsu.edu/yli/papers/WCMC10.pdf> (accessed August 1, 2014).
11. Grusho, A. A., N. A. Grusho, and E. E. Timonina. 2009. Metody zashchity informatsii ot atak s pomoshch'yu skrytykh kanalov i vrazhdebnnykh programmno-apparatnykh agentov v raspredelennykh sistemakh. [Information security methods from attacks by means of the covert channels and hostile hardware-software agents in distributed systems]. *Vestnik RGGU. Ser. "Informatika. Zashchita informatsii. Matematika"* [RGGU Bulletin. Informatics. Information security. Mathematician ser.] 10:33–45.
12. Wireless LAN. *Data Network Resource*. Available at: <http://www.rhyshaden.com/wireless.htm> (accessed August 1, 2014).
13. Microsoft. 2014. *Data flow in the OSI model*. Available at: <http://technet.microsoft.com/en-us/library/cc977591.aspx> (accessed August 1, 2014).
14. Tamosoft. Available at: <http://www.tamos.ru/products/commview/> (accessed October 2, 2014).

Received August 1, 2014

## Contributor

**Grusho Nick A.** (b. 1982) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; info@itake.ru

## УМНОЖИТЕЛЬ С НАКОПЛЕНИЕМ: МЕТОДОЛОГИЧЕСКИЕ АСПЕКТЫ\*

*И. А. Соколов<sup>1</sup>, Ю. А. Степченков<sup>2</sup>, С. Г. Бобков<sup>3</sup>, Ю. В. Рождественский<sup>4</sup>,  
Ю. Г. Дьяченко<sup>5</sup>*

**Аннотация:** Представлены подходы к проектированию самосинхронной (СС) аппаратуры и рассмотрены условия внутрисистемной интеграции синхронных и СС-устройств в рамках супер-ЭВМ на примере разработки СС-устройства умножения-сложения, соответствующего стандарту IEEE 754 и выполняющего либо одну операцию двойной точности, либо одновременно две операции одинарной точности над тремя операндами. Устройство реализовано по технологии КМОП (комплементарная структура металл-оксид-полупроводник) с проектными нормами 65 нм с использованием в умножителе СС троичного кодирования. В зависимости от исполнения оно работает с асинхронным или синхронным окружением и обеспечивает производительность на уровне не менее 1 гигафлопса при времени задержки формирования результата относительно поступления входных operandов не более 6 нс.

**Ключевые слова:** самосинхронная схема; супер-ЭВМ; умножитель с накоплением; сумматор; конвейер; энергоэффективность

**DOI:** 10.14357/08696527140304

### 1 Введение

Современные супер-ЭВМ петафлопсного класса используют менее 35% вычислительной мощности для непосредственных вычислений [1]. Остальные ресурсы расходуются на обеспечение достоверности результатов вычислений. При этом их наработка на отказ составляет  $\sim 55$  ч (данные на 2013 г. соответствуют супер-ЭВМ, содержащим 100 000 ядер вычислительных устройств). Для супер-ЭВМ эксафлопсного класса, где число ядер повышается почти на три порядка, а нормы проектирования сверхбольших интегральных схем (СБИС) сокращаются в 2–3 раза, показатели надежности вычислений и наработка на отказ могут оказаться существенно хуже. Решение этой проблемы возможно либо кардинальной заменой технологического базиса вычислителей, либо использова-

\* Исследование выполнено при финансовой поддержке РФФИ (проект 13-07-12068 офи\_м).

<sup>1</sup> Институт проблем информатики Российской академии наук, ISokolov@ipiran.ru

<sup>2</sup> Институт проблем информатики Российской академии наук, YStepchenkov@ipiran.ru

<sup>3</sup> Научно-исследовательский институт системных исследований Российской академии наук, bobkov@cs.niisi.ras.ru

<sup>4</sup> Институт проблем информатики Российской академии наук, YRogdest@ipiran.ru

<sup>5</sup> Институт проблем информатики Российской академии наук, diaura@mail.ru

нием аппаратных методов контроля надежности и саморемонта вычислительных средств в темпе реальных вычислений. Создание и глубокая технологическая проработка нового базиса вычислительных систем (ВС) видится в весьма отдаленной перспективе, в то время как для второго варианта решения можно уже сейчас предложить СС-схемотехнику, достаточно хорошо проработанную и отвечающую выдвигаемым требованиям.

Данная работа посвящена расширению класса СС-блоков вычислительного ядра для супер-ЭВМ. Ранее был разработан и изготовлен квазисамосинхронный блок деления/извлечения корня [3, 2] 64/32-разрядной арифметики в технологии 180 и 65 нм. Здесь излагаются основные принципы проектирования СС-устройства умножения-сложения 64/32-разрядности гигафлопсного класса для супер-ЭВМ, учитывающие особенности функционирования СС-схем.

Несмотря на теоретически доказанные и практически подтвержденные преимущества СС-изделий (в том числе разработанных, изготовленных и испытанных авторами настоящей статьи), коммерческих СС-изделий немного [4], а сведения об их успешной эксплуатации в рамках крупных ВС отсутствуют. Представляется, что одной из причин является проблема встраивания СС-изделий в глобальное синхронное окружение. В ряде работ В. И. Варшавского была предложена методология построения крупных ВС в виде GALA-систем (Global Asynchronous / Local Arbitrary). Если бы ее приняли на вооружение конструкторы супер-ЭВМ, то проблема внедрения СС-изделий была бы решена. На уровне системного интерфейса супер-ЭВМ реализуется асинхронный интерфейс обмена взаимодействующих устройств, а на уровне локального интерфейса — произвольный. Таким образом, на уровне локального интерфейса проектировщик может использовать широкую гамму решений — от реализации полной самосинхронизации до использования локального генератора. Все зависит от цели проектируемого изделия.

Но именно при такой стратегической линии построения супер-ЭВМ обеспечивается возможность постепенного внедрения СС-изделий (по мере появления их на рынке) в общую структуру для достижения требуемых параметров (надежностных показателей или энергоэффективности).

Операция «умножение двух операндов и сложение произведения с третьим операндом» (Fused Multiply-Add, FMA) с 1990-х гг. стала широко внедряться в микропроцессоры общего и специального назначения. В виде одной инструкции она реализована в микропроцессорах IBM POWER1 (1990 г.) и старше, Fujitsu SPARC64 (1995 г.), HP PA-8000 (1996 г.), Intel Itanium (2001 г.), IBM Cell (2005 г.), в процессорах AMD (2012 г.), в графических процессорах NVIDIA Tesla T10 (GPGPU) [5]. В 2008 г. она вошла в стандарт IEEE 754 [6]. Эта операция обычно выполняется быстрее, чем пара последовательных инструкций умножения и сложения, и точнее за счет отсутствия промежуточного округления произведения. В большинстве публикаций рассматриваются синхронные устройства FMA [7, 8], но в последние годы все больше публикаций посвящается асинхронным FMA, например [9, 10]. Анализ последних показал, что они опира-

ются на использование так называемой спекулятивной индикации (Bounded delay) и слабых транзисторов. Использование спекулятивной индикации — окончание переходных процессов индицируется только на некотором критическом пути — позволяет резко сократить объем оборудования на индикаторную часть, но приводит к потери самопроверяемости относительно константных неисправностей. Использование слабых транзисторов позволяет уменьшить общее число транзисторов, но приводит к существенному снижению уровня помехоустойчивости аппаратуры.

Более того, использование указанных приемов уменьшает полезную зону работоспособности СС-изделий, особенно в области предельно низких питающих напряжений. Именно это преимущество СС-схем позволяет существенно снизить динамическую и статическую составляющие энергопотребления для той части аппаратуры супер-ЭВМ, которая в настоящий момент либо не используется вообще (не задействована в текущей задаче), либо может быть использована с существенно меньшей производительностью, путем динамического снижения питающего напряжения [11].

Указанные виды СС-схем являются неприемлемыми для использования в составе супер-ЭВМ. В рамках сложившейся на сегодня классификации целям супер-ЭВМ отвечают только два класса СС-схем:

- (1) схем, правильная работа которых не зависит от задержек элементов и от задержек в проводах при условии, что разница задержек в проводах после разветвления меньше минимально возможной задержки базового элемента библиотеки. Такие схемы называются независимыми от задержек элементов (НЗЭ). В западной классификации им соответствуют *speed-independend* (SI), или *quasi-delay-insensitive* схемы;
- (2) схем, правильная работа которых не зависит ни от задержек элементов, ни от задержек в соединяющих их проводах. По сложившейся западной классификации такие схемы называются *delay-insensitive* (DI) схемами.

Ярким представителем DI-схем является NCL-схема (Null Convention Logic) [12]. Проведенный авторами сравнительный схемотехнический анализ НЗЭ- и NCL-схем [13] показал, что «НЗЭ схемы, разрабатываемые в соответствии с методологией, продвигаемой ИПИ РАН, обладают меньшими аппаратурными затратами (в 4,49 раза при реализации двоичного счетчика, в 1,13 раза при реализации умножителя  $4 \times 4$ , до 2 раз при реализации более простых логических схем), большей производительностью и меньшим энергопотреблением по сравнению с NCL-схемами». Поэтому именно их целесообразно использовать в качестве схемотехнического базиса супер-ЭВМ эксафлопсного класса.

В отличие от синхронной, НЗЭ-схемотехника использует запрос-ответную дисциплину взаимодействия блоков, СС-кодирование информационных сигналов и развитую систему индикации развития вычислительного процесса. Поэтому появление любой константной неисправности в любой цепи такой схемы приводит к остановке вычислений на соответствующем индикаторе и немедленной локали-

зации проблемы. Это позволяет практически в темпе вычислений предоставить резервный аппаратный ресурс и продолжить безошибочную работу всего устройства (осуществляется саморемонт). В ряде случаев применение НЗЭ-схемотехники может позволить в несколько раз увеличить реальную производительность супер-ЭВМ эксафлопсного класса.

Независимые от задержек элементов схемы обеспечивают уменьшение энергетических расходов, связанных с изъятием из схемы генератора тактовых импульсов и «клокового дерева», которые определяют значительную долю динамической и статической составляющих рассеиваемой мощности — от 30% до 50% от общего энергопотребления синхронной схемы. Исполнение аппаратуры в НЗЭ-базисе обеспечивает автоматический перевод в энергосберегающий режим той части аппаратуры, которая не используется в текущем цикле обработки информации. Сохранение работоспособности НЗЭ-схем при сверхмальных значениях питающих напряжений открывает широкие перспективы для проектирования энергоэффективных изделий. Введение программно-аппаратных возможностей регулировки уровня питающих напряжений позволит существенно снизить энергопотребление микросхем в ряде режимов.

Платой за такие преимущества является аппаратная избыточность и дополнительные временные затраты на индикацию и дополнительную фазу «гашения» в работе СС-элементов. Грамотное проектирование этих схем позволяет существенно снизить эту избыточность, а в ряде случаев [14] и получить лучшие, чем в синхронных схемах, результаты.

Цель данной статьи — оценка подходов к разработке SI-устройства умножения-сложения (SIFMA) по стандарту IEEE 754 [6], обладающего расширенными возможностями и ориентированного на использование в супер-ЭВМ.

## 2 Особенности разрабатываемого SIFMA

Область применения данного SIFMA выдвигает в качестве основного требования минимальное энергопотребление при достаточно высокой производительности [1]. Предполагаемая производительность — 1–4 Гфлопс. Она определяется сравнительно невысокой тактовой частотой для машин эксафлопсного класса и большим числом узлов SIFMA на одну СБИС. Анализ минимаксной кривой, построенной в координатах энергопотребление–площадь в соответствии с методикой, приведенной в [15], применительно к технологии 65 нм и к производительности SIFMA, позволил определить основные характеристики прототипа, выбрать структурную схему SIFMA и определить число ступеней в конвейерной реализации на 1 Гфлопс равным 8. Предлагаемый вариант SIFMA разрабатывался в рамках стандартной технологии 65 нм с ограничением латентности операций в 6 нс.

SIFMA предназначен для потоковой обработки троек operandов путем выполнения умножения двух operandов и сложения с третьим operandом или вычитания из третьего операнда.

Операнды двойной точности представляются как 64-разрядные числа, а одинарные — как 32-разрядные числа. Это позволяет использовать 64-разрядные операнды для передачи сразу двух 32-разрядных чисел. Тем самым создается предпосылка для опциональной обработки одним устройством не только операндов двойной точности, но и одновременной обработки двух однотипных операций одинарной точности.

Разработанный SIFMA выполняет операции в соответствии со стандартом IEEE 754 со следующими уточнениями:

- на вход поступают нормализованные операнды;
- перемножаемые операнды не являются нулевыми;
- тройка операндов содержит либо три числа двойной точности, либо шесть чисел одинарной точности; в последнем случае одновременно выполняются две операции над тройками операндов одинарной точности.

Кроме того, SIFMA опционально выполняет сразу две операции: сложение произведения двух операндов с третьим операндом и/или вычитание произведения из третьего операнда. Рассматриваемое устройство SIFMA реализует следующую совокупность действий по обработке входных операндов [7]:

1. Вычисление экспоненты результата.
2. Умножение мантисс первых двух операндов.
3. Выравнивание мантиссы третьего операнда.
4. Инвертирование выравненной мантиссы третьего операнда в случае фактического вычитания произведения из третьего операнда или наоборот.
5. Сложение и/или вычитание мантисс произведения и третьего операнда.
6. Получение модуля результата при его отрицательном значении.
7. Нормализация модуля суммы и/или разности.
8. Округление нормализованного результата.
9. Постнормализация.

Главной особенностью описываемого SIFMA является его принадлежность к классу устройств, правильно работающих при любых задержках элементов, на которых они реализованы, — к классу устройств, не зависящих от задержек элементов [3, 2]. Время выполнения операции на нем определяется не тактовым сигналом, а условиями эксплуатации: напряжением питания, температурой окружающей среды — и характером обрабатываемых операндов.

Достигаются такие характеристики с помощью кодирования сигналов специальными СС-кодами и средств индикации завершения отдельных этапов обработки данных. В подавляющем большинстве НЭЭ-схем используется парафазный код и индикаторы на Г-триггерах.

Фундаментальной проблемой проектирования СС-схем для супер-ЭВМ является необходимость индикации завершения каждого этапа вычислений для

всех разрядов машинного слова перед началом последующего этапа обработки. Такие индикаторы требуют существенных аппаратных затрат и привносят дополнительную задержку в работу процессора ЭВМ вследствие большой длины машинного слова и избыточности СС-кодирования.

Существует целый ряд возможных решений этой проблемы:

- использование специальных многовходовых индикаторных элементов;
- сокращение числа индицируемых разрядов;
- максимально возможное сокращение числа вычислительных элементов, требующих аппаратной индикации завершения процесса вычислений.

Попытки создания эффективных многовходовых индикаторных элементов предпринимались многократно [3, 6, 7]. В ряде случаев удавалось получить некоторое сокращение временных затрат, как правило за счет снижения помехоустойчивости вследствие использования слаботочных активных элементов.

Сокращение числа индицируемых разрядов возможно либо за счет отказа от строгой самосинхронности и индикации завершения процесса по неполному набору разрядов машинного слова [1, 5], либо за счет перехода к двуполярному питанию, либо за счет использования  $n$ -значной арифметики (при  $n > 2$ ) во внутреннем представлении чисел [16].

Первое решение существенно снижает затраты на индикацию, но в той же мере ликвидирует все преимущества СС-схем. Второй вариант уменьшает в 2 раза число индицируемых сигналов, но ведет к усложнению аппаратуры. Третий вариант приводит к усложненным алгоритмам непосредственной математической обработки одного разряда. Положительного эффекта в этом случае можно достичь лишь за счет существенного сокращения общей длительности вычислительного процесса с минимизацией числа разрядов, требующих индикации.

Применительно к SIFMA это требование означает минимизацию и, если это возможно, отказ от индикации любых промежуточных стадий процесса умножения двух чисел большой разрядности до его завершения.

## 2.1 Особенности реализации SIFMA большой разрядности

Наиболее эффективные алгоритмы умножения, применяемые практически во всех самых мощных и современных процессорах, базируются на алгоритмах кодирования Бута и сложения на базе дерева Уоллеса и их модификациях. При этом основные временные и аппаратные затраты приходятся на процедуру сложения частичных произведений, формируемым алгоритмом Бута, и реализуемую с помощью дерева Уоллеса. Борьба за повышение быстродействия привела к наращиванию конвейеризации в реализации алгоритма Уоллеса.

Временные затраты на операцию умножения в конвейерной реализации  $t_{\text{умн.к}}$  можно ориентировочно оценить как

$$t_{\text{умн.к}} = t_{\text{конв.н}} + \frac{t_{\text{умн}}}{n},$$

где  $t_{\text{конв.н}}$  — дополнительные временные затраты на организацию конвейерной обработки данных на каждом этапе конвейера;  $t_{\text{умн}}$  — чистое время выполнения операции умножения без использования конвейеризации;  $n$  — число ступеней конвейера для реализации умножения.

В синхронных схемах, не требующих средств индикации завершения процессов вычислений на каждом этапе конвейера, удавалось реализовать условие

$$t_{\text{умн}} \gg t_{\text{конв.н}}.$$

Это позволяло использовать конвейеризацию в умножителе с  $n = 5\text{--}7$  и выше. Однако при этом существенно росли аппаратные затраты и энергопотребление.

Переход в область субмикронных технологий (ограничение плотности тока на кристалле) и появление многоядерных процессоров внесли свои корректизы в этот процесс. В современных многоядерных процессорах стараются либо снизить до минимума число ступеней конвейера в блоке реализации дерева Уоллеса, либо отказаться от них вообще. Эта тенденция вполне соответствует и SIFMA-реализации. Вследствие наличия индикаторных цепей для них

$$t_{\text{конв.н}} \sim 0,5t_{\text{умн}}. \quad (1)$$

Это снижает эффективность конвейеризации или заставляет от нее отказаться.

Вторая проблема СС-реализации умножителя — наличие фазы гашения в двухфазной СС-дисциплине. В случае SIFMA эта проблема осложняется тем, что схема является полностью комбинаторной. Переход схемы в спейсерное состояние проходит последовательно. Все каскады дерева Уоллеса переключаются аналогично отработке рабочего состояния. Это делает длительность переключения и энергетические затраты фазы гашения практически равными соответствующим параметрам рабочей фазы. Решение этой проблемы зависит от жесткости требований к быстродействию умножителя и предполагает два варианта:

- (1) сокращение длительности фазы гашения за счет дополнительных цепей быстрой инициализации информационных структур умножителя;
- (2) использование двух параллельно и противофазно работающих умножителей.

Введение дополнительных цепей быстрой инициализации существенно сокращает временные затраты на переход в спейсерное состояние информационных цепей, и длительность фазы гашения в наибольшей степени будет определяться работой индикаторных схем. Однако соотношение (1) указывает на ограниченность такого сокращения. Кроме того, дополнительные цепи инициализации увеличивают энергопотребление умножителя и привносят дополнительную задержку в основной рабочий цикл. Эффективность сокращения временных затрат в этом случае оказывается не более 20%–30%.

Использование двух параллельно работающих умножителей в 2 раза увеличивает производительность схемы при двукратном увеличении аппаратных

затрат. Однако такое решение более предпочтительно, так как обеспечивает наименьшее энергопотребление на одну операцию из-за отсутствия дополнительных цепей инициализации.

## 2.2 Сокращение числа вычислительных этапов в работе умножителя

Этот подход является наиболее эффективным с точки зрения критерия быстродействие–энергопотребление, но одновременно и наиболее сложным в реализации. Он требует большого объема исследований различных алгоритмов умножения применительно именно к СС-решению с его дополнительными требованиями. Поскольку блок умножения практически является основным блоком всех процессоров, его разработке было уделено самое серьезное внимание с момента создания вычислительных систем и предложить что-то принципиально новое не представляется возможным. В качестве прототипа было выбрано алгоритмическое решение блока, реализующего дерево Уоллеса для синхронных схем, изложенное в работе [2]. Оно позволило создать самый быстрый блок умножения к началу текущего столетия и определило основные архитектурные направления в выпускаемых процессорах большой вычислительной мощности всеми ведущими мировыми фирмами. Суть предложенного авторами метода заключается в применении избыточного кодирования операндов дерева Уоллеса. Каждый двоичный разряд операнда представляется в виде двух двоичных разрядов согласно приведенной таблице.

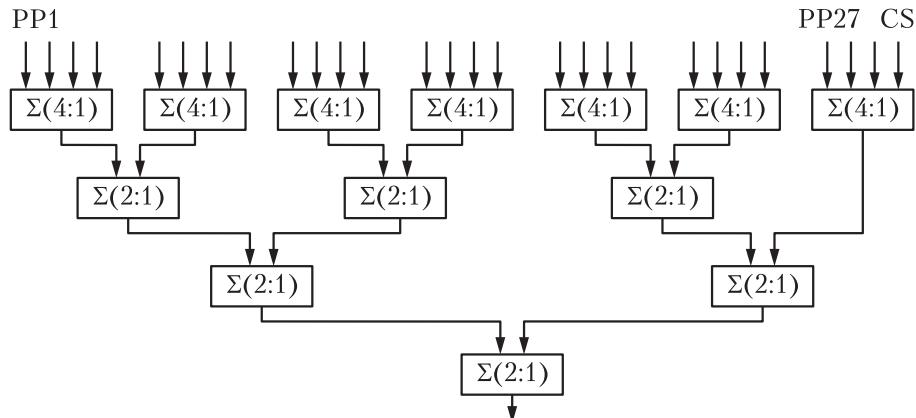
Такая избыточность позволяет описывать сумму двух однобитных чисел в прямом и обратном кодах одним кодовым числом и производить операцию сложения в одном кодовом разряде с сохранением переноса. Эта схема сложения обеспечивает коэффициент сжатия в дереве Уоллеса 2 : 1, реализуя наиболее быстрое сложение практически без аппаратной избыточности.

Попытка использовать парафазное кодирование для каждого из двоичных разрядов этого представления, как это обычно делается при переходе от синхронных схем к СС-решениям, приводит к заметному снижению эффективности алгоритма и большой аппаратной избыточности.

Предложенная схема кодирования очень похожа на парафазное СС-кодирование двоичного разряда и вызывает желание задействовать третье неиспользуемое состояние парафазного кода для описания «–1». Однако эта кажущаяся по-

Сравнение методов кодирования

Синхронное двоичное кодирование			Самосинхронное троичное кодирование			
Кодируемое состояние	Двоичный код		Кодируемое состояние	Троичный код		
	A	B		Ap	Am	An
+1	1	0	+1	1	0	0
0	0	0	0	0	0	1
–1	0	1	–1	0	1	0
Не используется	1	1	Спейсер	0	0	0



**Рис. 1** Схема СС реализации дерева Уоллеса для 53-разрядных чисел для троичного кодирования двоичных разрядов

хожесть приводит к несамосинхронному коду представления двоичного разряда. Согласно формуле построения кодов двухфазной СС-дисциплины операций для описания трех состояний и спейсера минимально необходимы три бита [4]. После проведенного анализа возможных СС-кодов авторами была предложена схема кодирования, представленная в таблице. Выбранный код оказался очень удачным.

Полученная СС-реализация дерева Уоллеса для сложения операндов, кодированных алгоритмом Бута, для 53-разрядных чисел представлена на рис. 1. Здесь PP1–PP27 — частичные произведения в парафазном коде, CS — корректирующее слагаемое в парафазном коде, а остальные внутренние и выходные сигналы представлены в СС троичном коде. За счет СС троичного кодирования внутренних сигналов, предложенной в [13] методики преобразования каждого второго частичного произведения и учета дополнительного корректирующего частичного произведения CS первый каскад дерева Уоллеса обеспечивает сжатие входных operandов с 27 частичных произведений до 7 троичных сигналов.

Для сравнения на рис. 2 показана соответствующая СС-реализация дерева Уоллеса в парафазном коде. Здесь CI1–CI24 — входные переносы из предыдущего разряда дерева, CO1–CO24 — выходные переносы в следующий разряд. Все сигналы парафазные. Троичная СС-реализация обеспечивает сокращение временных затрат по сравнению с классическим СС алгоритмом более чем на 20%. Дополнительное снижение аппаратных затрат также составило около 20%, а число этапов сжатия снизилось с 7 до 4.

### 3 Проблемы индикации SIFMA

Корректная работа любой SI-схемы обеспечивается индицированием окончания всех переключений схемы. Схемы, в которых индицируются выходы всех

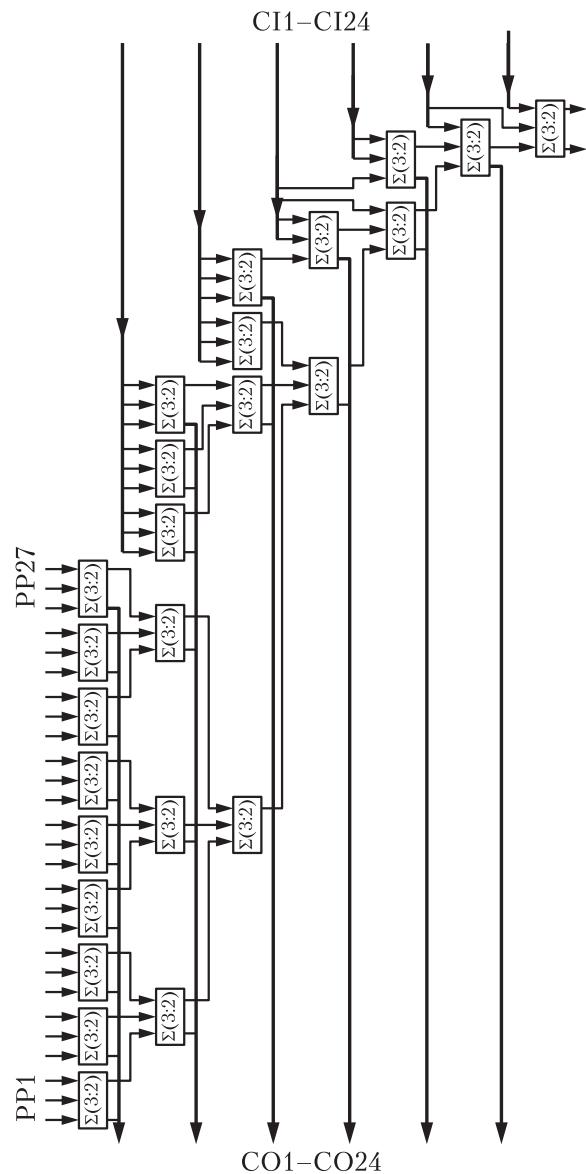


Рис. 2 Дерево Уоллеса в параллельном коде

без исключения элементов в обеих фазах работы (спейсерной и рабочей), называются полумодулярными [17]. Однако индикаторная подсхема, как правило, существенно замедляет работу полумодулярной многоразрядной схемы. «Узким горлом» многоразрядных СС-устройств является общий индикатор устройства, объединяющий все поразрядные индикаторные сигналы в один общий индикаторный сигнал.

Особенностью предлагаемого подхода к реализации SIFMA являются следующие принципы организации индикации:

- необходимая и достаточная, но упрощенная индикация рабочей фазы каждой ступени конвейера;
- использование спекулятивной индикации для ускорения взаимодействия ступеней конвейера.

С учетом того, что результат обработки данных комбинационной схемой (КС) в каждой ступени конвейера сохраняется в выходном регистре ступени и там индицируется в обеих фазах работы, нет необходимости индицировать КС в рабочей фазе. Используемое парафазное и специальное троичное СС-кодирование данных на всех этапах обработки гарантирует однократное переключение выходов КС из спейсера в рабочее состояние при переключении всей ступени конвейера из спейсера в рабочую фазу. Поэтому появление рабочего состояния после спейсера на информационных входах всех разрядов выходного регистра ступени гарантирует готовность результата.

Напротив, при переключении КС в спейсер индицируются все элементы, так как перед переходом в следующее рабочее состояние должна быть уверенность в том, что все элементы схемы переключились в спейсер. В противном случае возможны «гонки» и неоднократные срабатывания выходов элементов, что является нарушением принципов СС-реализации.

Поскольку КС в каждой ступени конвейера SIFMA является сложным функциональным многоразрядным устройством, такая упрощенная индикация существенно сокращает аппаратурные затраты и ускоряет формирование индикаторных сигналов, участвующих в запрос-ответном взаимодействии ступеней конвейера.

Спекулятивный индикатор отслеживает окончание переключения только элементов, стоящих на критическом пути обработки данных ступенью конвейера. Фактически это индикатор одного разряда тракта обработки многоразрядных данных. В многоразрядной схеме он формируется с гораздо меньшей задержкой, чем общий индикатор. Поскольку взаимодействие ступеней конвейера SI-схемы основано на двухфазной дисциплине работы [17], спекулятивный индикатор текущей ступени может использоваться в качестве сигнала разрешения переключения в противоположную фазу работы предшествующей ступени. Это позволяет предыдущей ступени начать переключение в очередную фазу, не дожидаясь подтверждения окончания переключения всех разрядов текущей ступени,

что существенно (на 30%–50% в зависимости от сложности схемы и разрядности данных) ускоряет работу конвейера.

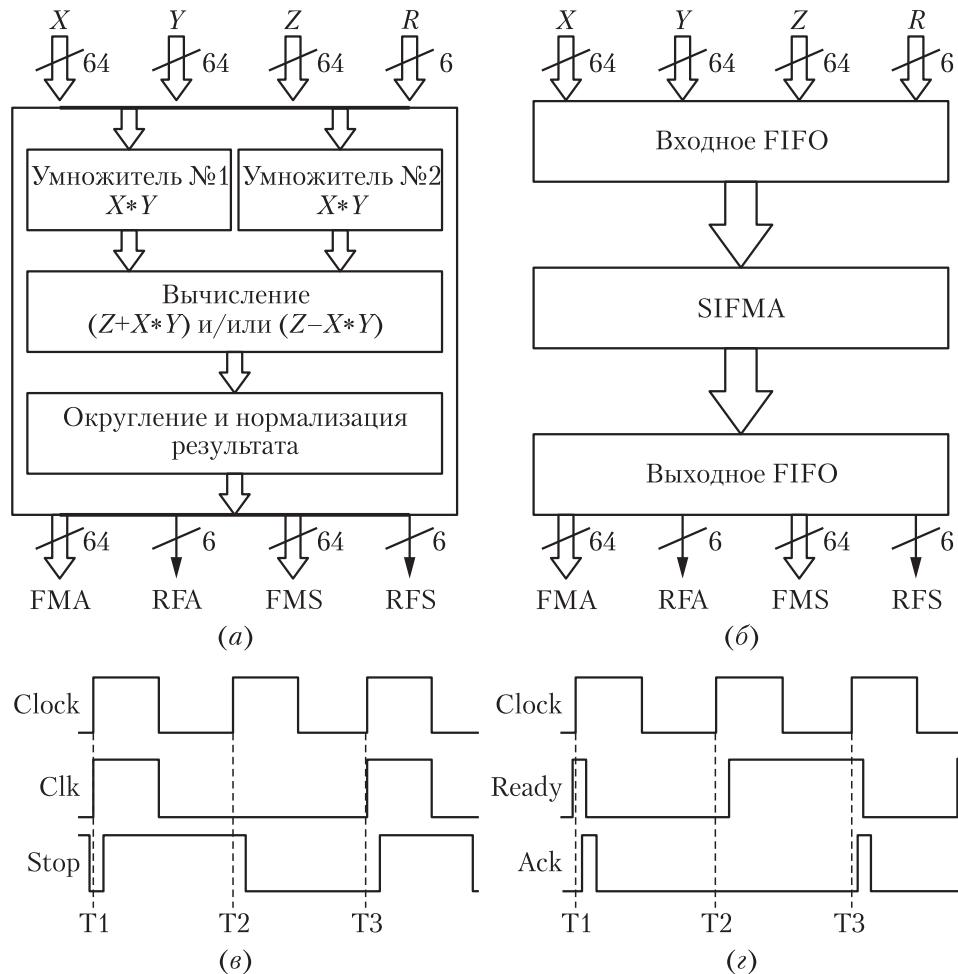
Логическим обоснованием такого подхода к индикации многоразрядных SI-схем служит следующее. Спекулятивный индикатор контролирует один из аналогичных разрядов тракта обработки данных, в том числе разряд выходного регистра. При правильном выборе соответствующего разряда в подавляющем большинстве случаев он будет переключаться не раньше (а если и раньше, то не намного) остальных разрядов схемы. Поэтому спекулятивный индикатор как бы отражает работу всех разрядов ступени. Общий индикатор оказывается более задержанным из-за того, что он «собирает» индикаторные сигналы со всех разрядов схемы с помощью «дерева» индикаторных элементов. Чем больше разрядов в схеме, тем более задержанным окажется общий индикаторный сигнал в сравнении со спекулятивным индикатором.

Стопроцентная самопроверяемость конвейера SIFMA обеспечивается использованием полного индикаторного выхода каждой ступени в качестве дополнительного разрешения переключения этой ступени конвейера в противоположную фазу работы. Такая схема по-прежнему индицирует константные неисправности в полном объеме, хотя уже и не обеспечивает бесперебойной работы при любых задержках элементов. С практической точки зрения она обладает всеми преимуществами СС-схем перед синхронными аналогами. С точки зрения полноты индикации рассматриваемый SIFMA не является полумодулярным, но при правильном выборе источника спекулятивного индикатора будет работоспособным в широком диапазоне условий эксплуатации (напряжения питания, температуры), технологических и электрических параметров компонентов и обязательно остановится при возникновении константной неисправности.

#### 4 Структурная схема SIFMA

Как основа для реализации супер-ЭВМ блок SIFMA должен обладать достаточным быстродействием. В синхронных схемах проблема быстродействия сложного вычислительного тракта наиболее эффективно решается с помощью конвейеризации вычислений. При этом разрядность схемы, если она достаточно регулярная, не играет роли.

В НЗЭ-схемах конвейеризация также используется. Но здесь уже число разрядов в схеме напрямую влияет на ее быстродействие, так как приводит к усложнению и замедлению подсхемы индикации. Кроме того, конвейеризация умножителя требует использования регистров для хранения промежуточных результатов, разрядность которых намного превышает разрядность входных операндов. Это дополнительно усложняет как саму схему, так и индикаторную подсхему и обеспечивает лишь небольшое увеличение быстродействия. Другими словами, повышение быстродействия самой сложной части SIFMA — умножителя — с помощью его конвейеризации оказывается «нерентабельным». Поэтому было принято решение для увеличения быстродействия SIFMA использовать



**Рис. 3** Умножитель с накоплением: (а) структурная схема; (б) интерфейс с окружением; (в) проблема входного интерфейса; (г) проблемы выходного интерфейса

два блока умножителя, работающих параллельно, как показано на структурной схеме (рис. 3, а). Здесь входные операнды — 64-разрядные обрабатываемые числа  $X$ ,  $Y$  и  $Z$  и 6-разрядные признаки операции  $R$  (тип округления, точность и характер операции). Тип округления — двухразрядный признак, определяющий способ округления результата в соответствии с требованиями стандарта IEEE 754. Характер операции — двухразрядный признак, задающий количество и тип операций, выполняемых над операндами. Использование признака характера операции позволяет минимизировать энергопотребление SIFMA в случаях,

когда требуется выполнение только одной операции (сложения или вычитания) над произведением и третьим операндом.

Отметим, что параллельно вычисляются только произведения операндов  $X$  и  $Y$  из двух последовательных троек операндов, поступающих на вход SIFMA.

Дальнейшие вычисления выполняются поочередно для каждой тройки входных операндов в одном общем тракте обработки данных. Такое решение обеспечивает наилучшее соотношение «быстродействие / аппаратурные затраты» при реализации SIFMA.

Время выполнения операции блоком SIFMA, как и любым НЗЭ-устройством, зависит от многих факторов: условий эксплуатации (напряжения питания, температуры окружающей среды), типа операции, вида операндов. Функциональные блоки SIFMA работают под управлением сигналов запрос-ответного взаимодействия и не требуют сигналов синхронизации.

Однако при работе с синхронным окружением возникает необходимость приема входных операндов по фронту сигнала системной частоты. В этом случае отсутствие потерь на приеме обеспечивается двумя способами:

- (1) подбором системной частоты исходя из наихудшего случая, при которой SIFMA заведомо успеет обработать любую тройку операндов при любых предусмотренных условиях эксплуатации,
- (2) использованием устройств сопряжения синхронных и асинхронных устройств, предотвращающих потерю операндов за счет приостановки «накачки» данных со стороны синхронного окружения.

Второй способ является более эффективным и быстродействующим, поскольку в его рамках легко реализовать буферизацию данных на входе/выходе SIFMA.

Для обеспечения максимального быстродействия SIFMA при работе с синхронным окружением схему на рис. 3, *a* предлагается дополнить входным и выходным НЗЭ-FIFO (first in, first out), как показано на рис. 3, *б*. Самосинхронный FIFO накапливает поступающие тройки операндов и признаки операции, обеспечивая формирование предупреждения о заполнении FIFO. Сигнал предупреждения Stop устанавливается ( $Stop = 1$ ) в том случае, если занята входная ячейка FIFO. Это происходит при приеме каждой тройки операндов, как показано на рис. 3, *в*, где Clock — сигнал глобальной синхронизации, а Clk — тактовый вход устройства SIFMA. С продвижением операндов вглубь FIFO по мере обработки предыдущих операндов входная ячейка освобождается и сигнал Stop снимается ( $Stop = 0$ ). Это дает синхронному окружению возможность вовремя приостановить подачу данных на вход SIFMA.

На рис. 3, *в* в моменты времени T1 и T3 SIFMA оказывается готовым к приему очередных данных и тактовый сигнал Clk формируется, а в момент времени T2 снятие сигнала Stop опаздывает по отношению к активному фронту тактового сигнала Clock и заставляет синхронное окружение приостановить запись данных в SIFMA.

Последующие блоки SIFMA последовательно обрабатывают поступающие данные в СС-режиме: по мере готовности данные передаются из текущего блока в следующий, а текущий блок переходит в режим ожидания данных из предыдущего блока.

Блок FIFO на выходе SIFMA также предотвращает потерю производительности SIFMA, когда время формирования очередного результата превысит период системной частоты, из-за чего один из циклов чтения результата синхронным окружением может быть пропущен, как показано на рис. 3, г.

Сигнал Ready индицирует готовность результата на выходе SIFMA. Синхронное окружение фиксирует его по активному фронту тактового сигнала Clock и сигналом Ack = 1 подтверждает прием данных. В моменты времени T1 и T3 результат оказывается готовым перед появлением очередного активного фронта Clock, а в момент времени T2 опаздывает по отношению к активному фронту тактового сигнала. Из-за этого возникает «холостой» такт: результат выполнения очередной операции с выхода SIFMA считывается с опозданием на один такт.

Использование входного и выходного НЗЭ-FIFO позволяет решить проблемы максимально быстрой «накачки» SIFMA входными operandами и в полной мере использовать способность SIFMA выполнять операцию умножения-сложения за разное время в зависимости от типа операции, вида operandов и условий эксплуатации.

В соответствии с описанными принципами были разработаны два варианта SIFMA [18] — для синхронного и асинхронного окружения. Они подтвердили требуемый уровень производительности и энергопотребления.

Изложенные результаты получены для версии кодирования алгоритма Бута Radix2, снижающей вдвое количество суммируемых operandов в дереве Уоллеса. Исследования других вариантов алгоритма Бута для Radix3 и Radix4 в синхронном исполнении [19] указывают на возможность получения дополнительного выигрыша в быстродействии и аппаратных затратах до 10%. Однако эти алгоритмы требуют предварительной подготовки operandов, равных трем значениям множимого, т. е. требуют изначально дополнительной операции сложения. Вопрос целесообразности использования таких алгоритмов в СС-схемах требует дополнительных исследований.

## 5 Заключение

В супер-ЭВМ эксафлопсного класса, когда число ядер достигает сотен миллионов, практически необходимо существенно усиливать аппаратную составляющую средств обеспечения надежности и достоверности результатов вычислений. Предложенная СС-схемотехника для реализации основных вычислительных узлов обеспечивает высокую эффективность в решении этой задачи.

Впервые в отечественной и зарубежной практике предпринята попытка разработки действительно 64/32-разрядного SIFMA-устройства в виде схемы,

поведение которой не зависит от задержек элементов и задержек в проводах до точки разветвления.

При разработке СС сумматоров и блоков умножения целесообразно использовать наряду с парафазным троичным СС-кодом.

Из всего многообразия алгоритмов следует выбирать аппаратно оправданные решения с минимальным числом этапов, требующих индикации завершения операций обработки.

Алгоритм суммирования по дереву Уоллеса рекомендуется по возможности целиком выполнять в одной ступени конвейера, применяя два параллельных вычислительных блока в противофазных режимах.

## Литература

1. Семенов Ю. А. Суперкомпьютеры и Watson. <http://book.itep.ru/10/supercomp.htm>.
2. Дьяченко Ю. Г., Рождественский Ю. В., Морозов Н. В., Степченков Д. Ю. Квазисамосинхронный вычислитель: практическая реализация // Проблемы разработки перспективных микро- и наноэлектронных систем: III Всеросс. науч.-технич. конф. (МЭС-2008): Сб. научных тр. — М.: ИППМ РАН, 2008. С. 435–440.
3. Степченков Ю. А., Дьяченко Ю. Г., Бобков С. Г. Квазисамосинхронный вычислитель: методологические и алгоритмические аспекты // Проблемы разработки перспективных микро- и наноэлектронных систем: III Всеросс. науч.-технич. конф. (МЭС-2008): Сб. научных тр. — М.: ИППМ РАН, 2008. С. 441–446.
4. Bink A., York R. ARM996HS: The first licensable, Clockless 32-bit processor core // IEEE Micro, 2007. Vol. 27. No. 2. P. 58–68.
5. Bing H. Acceleration of spiking neural network on general purpose graphics processor. Ph.D. Dissertation. — University of Dayton, 2010. 43 p.
6. IEEE Standard for Floating-Point Arithmetic 754-2008. — IEEE, 2008. doi:10.1109/IEEE STD. 2008.4610935.
7. Pillai R. V. K., Shah S. Y. A., Al-Khalili A. J., Al-Khalili D. Low power floating point MAFs — a comparative study // 6th Symposium (International) on Signal Processing and Its Applications, 2001. Vol. 1. P. 284–287.
8. Seidel P.-M. Multiple path IEEE floating-point fused multiply-add // 46th IEEE Midwest Symposium (International) on Circuits and Systems Proceedings, 2003. P. 1359–1362.
9. Noche J. R., Araneta J. C. An asynchronous IEEE floating-point arithmetic unit // Proc. Sci. Diliman, 2007. Vol. 19. No. 2. P. 12–22.
10. Manohar R., Sheikh B. R. Operand-optimized asynchronous floating-point units and method of use therefore. U.S. Patent No. 20130124592, May 2013.
11. Руткевич А., Бумагин А., Гондарь А. и др. Методы снижения энергопотребления в строго самосинхронных микропроцессорных схемах // Компоненты и технологии, 2009, № 9. С. 109–114.
12. Smith S. C., Di J. Designing asynchronous circuits using NULL Convention Logic (NCL) // Synthesis Lectures on Digital Circuits and Systems, 2009. Vol. 4. No. 1. P. 61–73.

13. Соколов И. А., Степченков Ю. А., Бобков С. Г., Захаров В. Н., Дьяченко Ю. Г., Рождественский Ю. В., Сурков А. В. Базис реализации супер-ЭВМ эксафлопсного класса // Информатика и её применения, 2014. Т. 8. Вып. 1. С. 45–70.
14. Степченков Ю. А., Петрухин В. С., Дьяченко Ю. Г. Опыт разработки самосинхронного ядра микроконтроллера на базовом матричном кристалле // Проблемы разработки перспективных микроэлектронных систем — 2005: Всеросс. науч.-технич. конф. (МЭС-2005): Сб. научных тр. — М.: ИППМ РАН, 2005. С. 235–242.
15. Galal S., Horowitz M. Energy-efficient floating-point unit design // IEEE Trans. Comput., 2011. Vol. 60. No. 7. P. 913–922.
16. Бриллюэн Л. Наука и теория информации / Пер с англ. — М.: Физматлит, 1960. 391 с. (Brillouin L. Science and information theory. New York: Academic Press Publs., 1956. 350 р.)
17. Автоматное управление асинхронными процессами в ЭВМ и дискретных системах / Под ред. В. И. Варшавского. — М.: Наука, 1986. 400 с.
18. Степченков Ю. А., Рождественский Ю. В., Дьяченко Ю. Г., Морозов Н. В., Степченков Д. Ю. Рождественскене А. В., Сурков А. В. Самосинхронный умножитель с накоплением: практическая реализация // Системы и средства информатики, 2014. Т. 24. № 3. С. 63–77.
19. Bewick G. W. Fast multiplication: Algorithms and Implementation. Ph.D. Dissertation. — Stanford University, 1994. 155 p.

*Поступила в редакцию 20.08.14*

---

## FUSED MULTIPLY-ADD: METHODOLOGICAL ASPECTS

*I. Sokolov<sup>1</sup>, Y. Stepchenkov<sup>1</sup>, S. Bobkov<sup>2</sup>, Y. Rogdestvenski<sup>1</sup>, and Y. Diachenko<sup>1</sup>*

<sup>1</sup>Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333, 44-2 Vavilova Str., Russian Federation

<sup>2</sup>Scientific Research Institute for System Studies, Russian Academy of Sciences, 36 bld. 1 Nakhimovsky Prospekt, Moscow 117218, Russian Federation

**Abstract:** The paper presents approaches to designing self-timed (ST) equipment and analyzes conditions of in-system integration of synchronous and ST units in a supercomputer network taking the ST Fused Multiply-Add (FMA) unit as an example. Self-timed FMA complies with the IEEE 754 Standard and performs either one double precision FMA operation or one or two single precision operations simultaneously under three operands. It utilizes the ST-ternary encoding and the 65-nanometer CMOS (complementary metal–oxide–semiconductor) technology as the implementation basis. Depending on realization, it works with asynchronous or synchronous environment and provides not less than 1 GFlops performance with latency not more than 6 ns with respect to input data arrival.

**Keywords:** self-timed circuit; supercomputer; fused multiply-add; adder; pipeline; energy efficiency

**DOI:** 10.14357/08696527140304

## Acknowledgments

This project was financially supported by the Russian Foundation for Basic Research (project 13-07-12068 ofi\_m).

## References

1. Semenov, Yu. A. *Superkomp'yutery i Watson* [Supercomputers and Watson]. Available at: <http://book.itep.ru/10/supercomp.htm> (accessed August 15, 2014).
2. Diachenko, Y. G., Y. V. Rogdestvenski, N. V. Morozov, and D. Y. Stepchenkov. 2008. Kvazisamosinhronnyy vychislitel': Prakticheskaya realizatsiya [Quasi self-timed coprocessor: Practical implementation]. *Tr. Mezhdunar. konf. "Problemy Razrabotki Perspektivnykh Mikro- i Nanoelektronnykh Sistem"* [Problems of the Perspective Micro- and Nanoelectronic Systems Development Conference Proceedings]. Moscow. 435–440.
3. Stepchenkov, Y. A., Y. G. Diachenko, and S. G. Bobkov. 2008. Kvazisamosinhronnyy vychislitel': Metodologicheskie i algoritmicheskie aspekty [Quasi-self-timed co-processor: The methodological aspects]. *Tr. Mezhdunar. konf. "Problemy Razrabotki Perspektivnykh Mikro- i Nanoelektronnykh Sistem"* [Problems of the Perspective Micro- and Nanoelectronic Systems Development Conference Proceedings]. Moscow. 441–446.
4. Bink, A., and R. York. 2007. ARM996HS: The first licensable, Clockless 32-bit processor core. *IEEE Micro*. 27(2):58–68.
5. Bing, H. Acceleration of spiking neural network on general purpose graphics processor. Ph.D. Dissertation. University of Dayton, 2010. 43 p.
6. IEEE Computer Society. 2008. IEEE Standard for Floating-Point Arithmetic 754-2008. doi:10.1109/IEEE STD. 2008.4610935.
7. Pillai, R. V. K., S. Y. A. Shah, A. J. Al-Khalili, and D. Al-Khalili. 2001. Low power floating point MAFs—a comparative study. *6th Symposium (International) on Signal Processing and Its Applications Proceedings*. 1:284–287.
8. Seidel, P.-M. 2003. Multiple path IEEE floating-point fused multiply-add. *46th IEEE Midwest Symposium (International) on Circuits and Systems Proceedings*. 1359–1362.
9. Noche, J. R., and J. C. Araneta. 2007. An asynchronous IEEE floating-point arithmetic unit. *Proc. Sci. Diliman* 19(2):12–22.
10. Manohar, R., and B. R. Sheikh. May 2013. Operand-optimized asynchronous floating-point units and method of use therefore. U.S. Patent No. 20130124592.
11. Rutkevich, A., A. Bumagin, A. Gondar', et al. 2009. Metody snizheniya ener-gopotrebleniya v strogo samosinkhronnykh mikroprocessornykh skhemakh [Methods of decreasing energy consumption in strongly self-timed microprocessors]. *Komponenty i Tehnologii* [Components and Technologies] 9:109–114.
12. Smith, S. C., and J. Di. 2009. Designing asynchronous circuits using NULL Convention Logic (NCL). *Synthesis Lectures on Digital Circuits and Systems* 4(1):61–73.
13. Sokolov, I. A., Y. A. Stepchenkov, S. G. Bobkov, V. N. Zakharov, Y. G. Diachenko, Y. V. Rogdestvenski, and A. V. Surkov. 2014. Bazis realizatsii super-EVM eksaflop-snogo klassa [Implementation basis of ExaFlops class supercomputer]. *Informatika i ee Primenenie — Inform. Appl.* 8(1):45–70.

14. Stepchenkov, Y. A., V. S. Petruhin, and Y. G. Diachenko. 2005. Opyt razrabotki samosinkhronnogo yadra mikrokontrollera na bazovom matrichnom kristalle [The experience in microcontroller's self-timed core design on FPGA]. *Tr. Mezhdunar. konf. "Problemy Razrabotki Perspektivnykh Mikro- i Nanojelektronnykh Sistem"* [Problems of the Perspective Micro- and Nanoelectronic Systems Development Conference Proceedings]. Moscow. 235–242.
15. Galal, S., and M. Horowitz. 2011. Energy-efficient floating-point unit design. *IEEE Trans. Comput.* 60(7):913–922.
16. Brillouin, L. 1956. *Science and information theory*. New York: Academic Press Publs. 350 p.
17. Varshavsky, V. I., ed. 1990. *Self-timed control of concurrent processes: The design of aperiodic logical circuits in computers and discrete systems*. Mathematics and its applications ser. Dordrecht, The Netherlands: Kluwer Academic Publs. 418 p.
18. Stepchenkov, Y. A., Y. V. Rogdestvenski, Y. G. Diachenko, N. V. Morozov, and D. Y. Stepchenkov. 2014. Samosinkhronnyy umnozhitel' s nakopleniem: Prakticheskaya realizatsiya [Self-timed fused multiply-add unit: Practical implementation]. *Sistemy i Sredstva Informatiki — The Systems and Means of Informatics* 24(3):63–77.
19. Bewick, G. W. 1994. Fast multiplication: Algorithms and Implementation. Ph.D. Dissertation. Stanford University. 155 p.

Received August 20, 2014

## Contributors

**Sokolov Igor A.** (b. 1954) — Academician of the Russian Academy of Sciences, Doctor of Science in technology, Director, Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333, Russian Federation; ISokolov@ipiran.ru

**Stepchenkov Yuri A.** (b. 1951) — Candidate of Sciences (PhD) in technology, Head of Department, Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333, Russian Federation; YStepchenkov@ipiran.ru

**Bobkov Sergey G.** (b. 1955) — Doctor of Science in technology, Head of Department, Scientific Research Institute for System Studies, Russian Academy of Sciences, Moscow 117218, Russian Federation; Bobkov@cs.niisi.ras.ru

**Rogdestvenski Yuri V.** (b. 1952) — Candidate of Science (PhD) in technology, Head of Laboratory, Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333, Russian Federation; YRogdest@ipiran.ru

**Diachenko Yuri G.** (b. 1958) — Candidate of Science (PhD) in technology, senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333, Russian Federation; diaura@mail.ru

## САМОСИНХРОННЫЙ УМНОЖИТЕЛЬ С НАКОПЛЕНИЕМ: ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ\*

*Ю. А. Степченков<sup>1</sup>, Ю. Г. Дьяченко<sup>2</sup>, Ю. В. Рождественский<sup>3</sup>,  
Н. В. Морозов<sup>4</sup>, Д. Ю. Степченков<sup>5</sup>, А. В. Рождественскене<sup>6</sup>, А. В. Сурков<sup>7</sup>*

**Аннотация:** Статья посвящена результатам разработки вариантов независимого от задержек устройства умножения-сложения (SIFMA — Speed-Independent Fused Multiply-Add), соответствующего стандарту IEEE 754 и выполняющего либо одну операцию умножения с накоплением двойной точности, либо одну или две операции одинарной точности над тремя operandами. Устройство разработано по стандартной технологии КМОП (комплементарная структура металл–оксид–полупроводник) с проектными нормами 65 нм. Оно работает с синхронным или асинхронным окружением и обеспечивает среднюю производительность на уровне 1 Гфлопса при напряжении питания 1 В и температуре 25 °С. Энергопотребление при этом не превышает 970 мДж/ГГц.

**Ключевые слова:** самосинхронная схема; троичное кодирование; умножитель; сумматор; вычитатель; конвейер; индикация

**DOI:** 10.14357/08696527140305

### 1 Введение

Операция «умножение двух operandов и сложение с третьим operandом» (Fused Multiply-Add, FMA) — одна из наиболее часто используемых в распределенных вычислениях. В литературе известны многочисленные реализации данной операции, относящиеся к классу синхронных схем (см., например, [1–7]). Однако наилучшее сочетание потребительских характеристик блока FMA обеспечивается применением самосинхронных (CC) схем, не зависящих от задержек элементов (Speed Independent, SI).

\* Исследование выполнено при финансовой поддержке РФФИ (проект 13-07-12068 офф.).

<sup>1</sup> Институт проблем информатики Российской академии наук, YStepchenkov@ipiran.ru

<sup>2</sup> Институт проблем информатики Российской академии наук, diaura@mail.ru

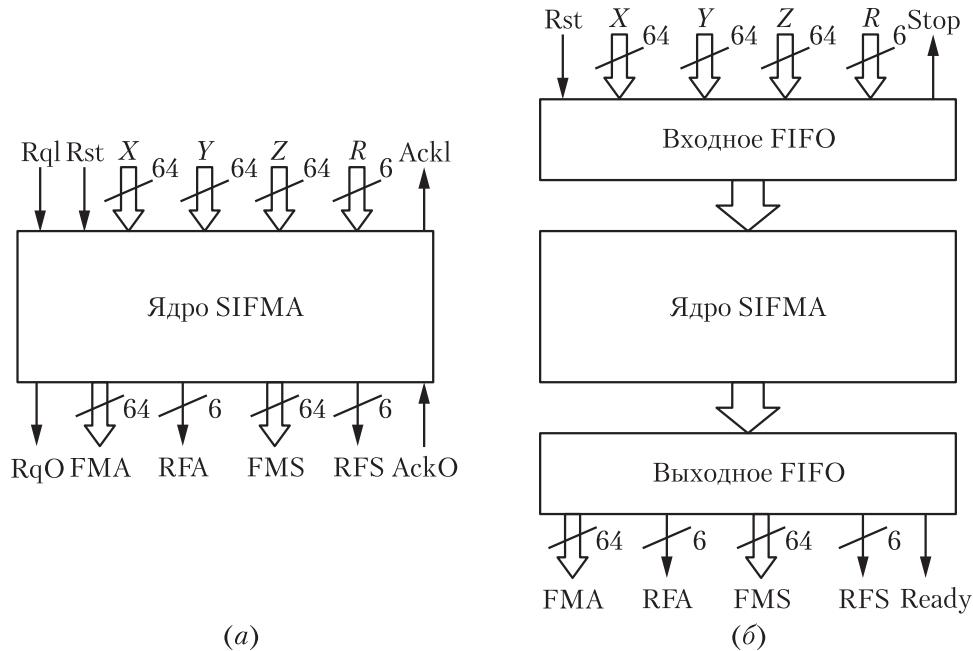
<sup>3</sup> Институт проблем информатики Российской академии наук, YRogdest@ipiran.ru

<sup>4</sup> Институт проблем информатики Российской академии наук, NMorozov@ipiran.ru

<sup>5</sup> Институт проблем информатики Российской академии наук, Stepchenkov@mail.ru

<sup>6</sup> Институт проблем информатики Российской академии наук, astarozd@yahoo.com

<sup>7</sup> Научно-исследовательский институт системных исследований Российской академии наук, surkov@cs.niisi.ras.ru



**Рис. 1** Варианты реализации SIFMA: с асинхронным (а) и синхронным (б) окружением

Цель данной статьи — разработка двух вариантов устройства умножения-сложения (SIFMA), предназначенных для работы с разным окружением. Оба варианта соответствуют стандарту IEEE 754 [8], принадлежат к классу SI-устройств [9], обладают расширенными функциональными возможностями и имеют сбалансированные характеристики. Первый вариант SIFMA (далее по тексту — асинхронный) предназначен для работы с асинхронным окружением и позволяет в максимальной степени использовать преимущества СС-схем (рис. 1, а). Второй вариант (синхронный) учитывает необходимость согласования входного и выходного интерфейса СС-устройства умножения-сложения с синхронным окружением (рис. 1, б).

## 2 Структурная схема SIFMA

Рассматриваемые варианты SIFMA разработаны на основе СС-ядра. Однако детерминированность процесса подачи операндов на вход SIFMA (на каждом такте системной частоты) при наличии синхронного окружения не позволяет в полной мере использовать тот факт, что время выполнения операции в SIFMA зависит от типа операции и значения операндов. Поэтому для обеспечения максимальной эффективности вычислительного процесса и достижения предель-

ного быстродействия вариант для синхронного окружения содержит устройства сопряжения с синхронным окружением — входное и выходное FIFO (first in, first out), также выполненные в стиле SI-устройств. Результирующая схема устройства показана на рис. 1, б.

Входные данные записываются в FIFO синхронно по фронту тактового сигнала Clk. Синхронное окружение не ждет подтверждения приема данных от SIFMA, но следит за дополнительным сигналом Stop, свидетельствующим о заполнении FIFO. Выходные данные такжечитываются синхронно по фронту тактового сигнала Clk, но только при наличии активного уровня на выходе Ready. В свою очередь, синхронное окружение информирует выходное FIFO сигналом Used о том, что оно приняло текущий результат и больше не нуждается в нем. В качестве такого сигнала может использоваться синхросигнал в регистре синхронного окружения, фиксирующем результат операции.

В качестве FIFO использованы CC полуплотные регистры сдвига [10, рис. 11.9] емкостью четыре слова данных. Схема одного FIFO показана на рис. 2. Отдельными оттенками изображены входная головка, внутренний элемент и выходная головка FIFO. За счет перекрестных обратных связей каждый разряд данного FIFO «проталкивает» принятый бит информации сразу до последней незанятой ячейки FIFO. Входная головка входного FIFO принимает от синхронного окружения информацию в традиционном для синхронных схем виде: данные плюс синхросигнал — и преобразует ее в парафазное представление.

Входная головка выходного FIFO работает с CC-устройствами, но ее схема идентична входной головке входного FIFO: роль синхросигнала играет поразрядный индикаторный выход предшествующего блока SIFMA.

Ядром обоих вариантов SIFMA является собственно вычислитель операции умножения-сложения, структурная схема которого показана на рис. 3. Его входной и выходной интерфейсы содержат все необходимые сигналы для организации взаимодействия SIFMA с асинхронным окружением.

Входные операнды — обрабатываемые числа  $X, Y, Z$  и признаки операции  $R$  (тип округления, точность и характер операции) поступают на вход SIFMA асинхронно. Их готовность подтверждается сигналом RqI со стороны источника. SIFMA запоминает данные во входных регистрах блоков умножителя и обработки экспонент (УОЭ1 и УОЭ2) и извещает об окончании этой процедуры сигналом AckI. Вход Rst служит для начального сброса устройства.

Отметим, что SIFMA способен одновременно выполнить две операции: « $Z + X * Y$ » и « $Z - X * Y$ » — благодаря наличию двух параллельных путей обработки произведения и третьего операнда.

На выходе SIFMA формируется сумма FMA и/или разность FMS и сопутствующие флаги результата RFA и RFS. Готовность результата индицируется сигналом RqO. Окончание чтения результата асинхронное окружение SIFMA подтверждает сигналом AckO.

Наиболее сложным функциональным блоком в составе SIFMA является умножитель мантисс  $53 \times 53$  (старший разряд — подразумеваемая, но отсутству-

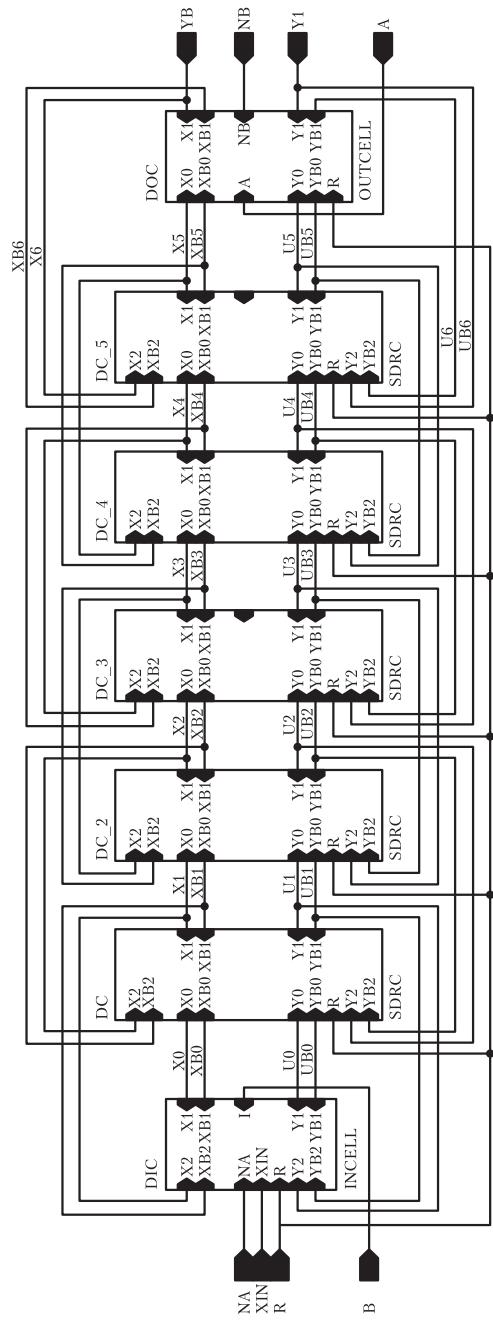
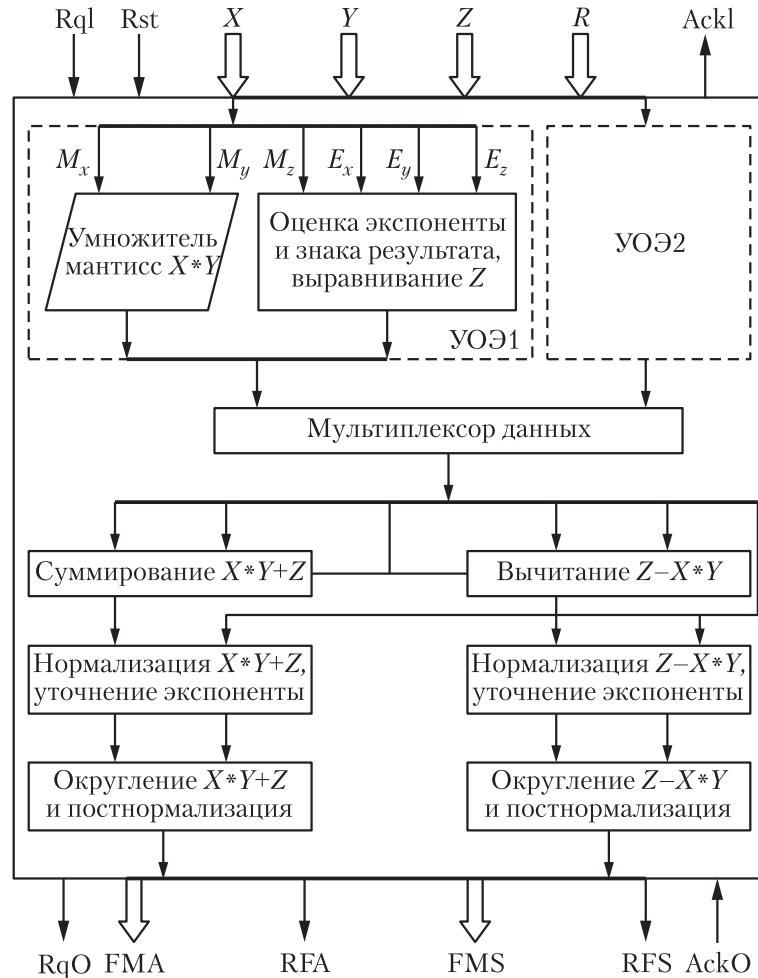


Рис. 2 Структурная схема FIFO

**Рис. 3** Структурная схема ядра SIFMA

ющая в формате представления чисел по стандарту IEEE 754 единица целых; младшие 52-разряда — мантисса числа). Он реализован на основе модифицированного алгоритма Бута (Booth) с основанием Radix-2 и дерева Уоллеса (Wallace) [11, 12]. Выходом всего дерева Уоллеса являются два 106-разрядных операнда: суммы и межразрядного переноса. Двоичный результат перемножения двух входных операндов получается с помощью их сложения на 106-разрядном сумматоре с распространением переноса.

Исследования показали, что повышение быстродействия при практических затратах достигается при использовании специального

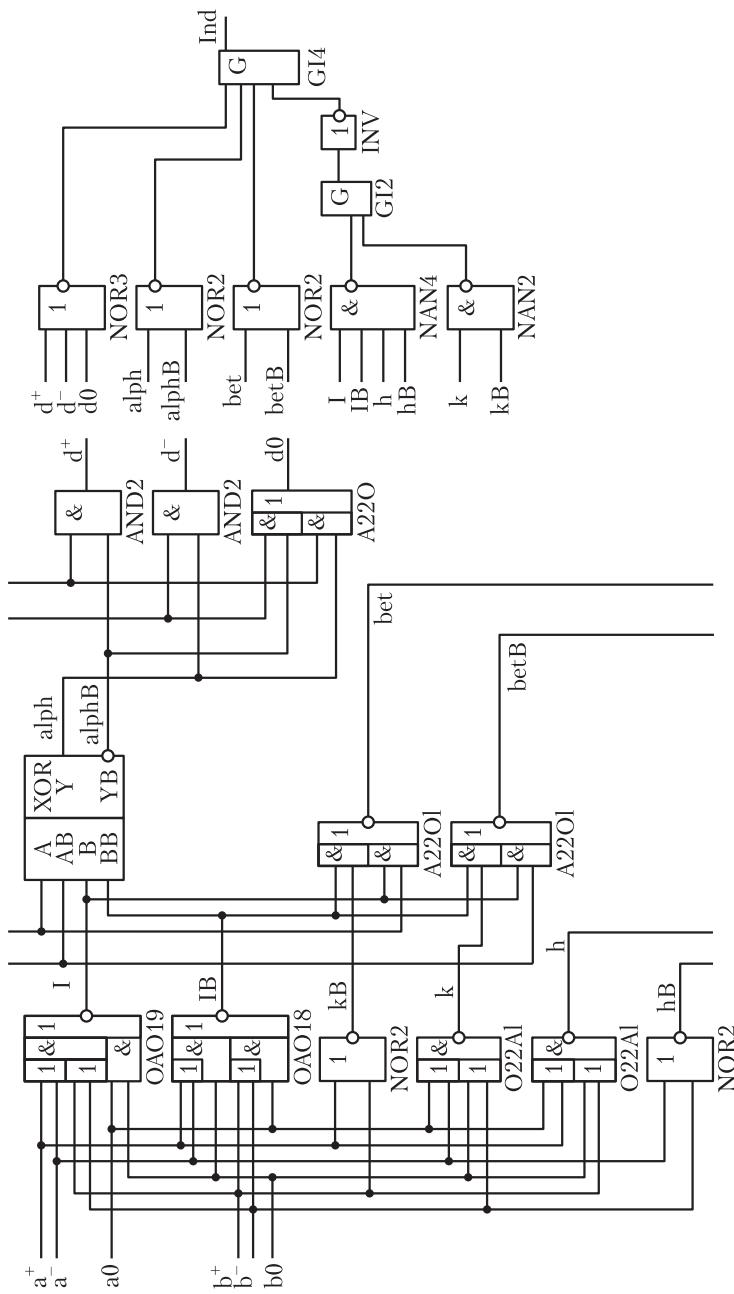


Рис. 4 Самосинхронный триномиальный сумматор

СС-кодирования, основанного на троичном избыточном представлении обрабатываемых операндов. Схема одного разряда сумматора таких операндов показана на рис. 4.

Специальное форматирование входных данных для алгоритма Бута позволяет в рамках одного умножения получить либо один результат операции двойной точности, либо сразу два результата двух операций одинарной точности.

Предлагаемая реализация SIFMA включает два блока УОЭ, работающих параллельно. Это обеспечивает максимальное быстродействие и сбалансированность ступеней конвейера при разумных аппаратурных затратах. Каждый блок УОЭ выполняет умножение операндов  $X$  и  $Y$ , а также анализ и обработку экспонент всех трех операндов и выравнивание операнда  $Z$ .

Очередность предоставления новых операндов блокам УОЭ не зависит от времени вычисления каждого блока и определяется простым чередованием. Это позволяет отказаться от использования арбитража как на входе блоков УОЭ, так и на их выходе и сохранить последовательность появления результата на выходе SIFMA, соответствующей порядку задания операндов на входе. Таким образом, окружение SIFMA всегда знает, результат какой именно операции присутствует на его выходе: SIFMA представляет собой как бы одно FIFO со сложной функциональной начинкой.

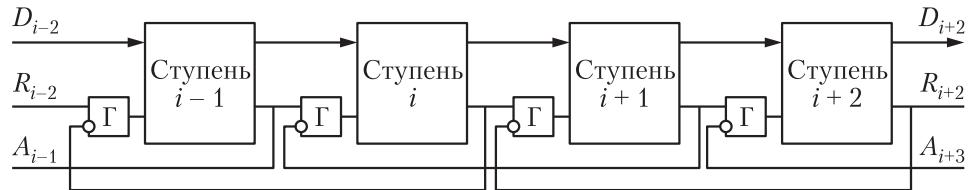
Последующие блоки ядра SIFMA обрабатывают поступающие данные в СС-режиме: по мере готовности данные передаются из текущего блока в следующий. Причем мультиплексор данных на выходе блоков УОЭ выбирает произведение мантисс двух операндов, выравненную в соответствии со значением экспонент всех трех операндов мантиссу третьего операнда и экспоненту результата на выходе активного в данный момент канала и распараллеливает дальнейшие вычисления. Сумма и разность выравненного третьего операнда и произведения первых двух операндов вычисляются одновременно для ускорения работы SIFMA.

Результаты выполнения обеих операций записываются в выходное FIFO. Операция считается выполненной, если результаты и сложения, и вычитания зафиксированы в выходном FIFO. Максимальная производительность в СС-устройствах, как и в синхронных схемах, обеспечивается конвейерной архитектурой.

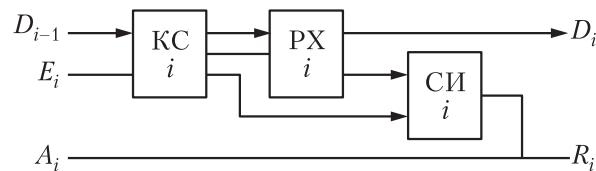
### 3 Конвейер SIFMA

Конвейер СС-устройства обычно строится на основе запрос-ответного взаимодействия между ступенями (рис. 5). Гистерезисные триггеры ( $\Gamma$ -триггеры [10]) на основе индикаторных выходов предыдущей и последующей ступени конвейера формируют сигналы управления, разрешающие переключение соответствующей ступени из рабочей фазы в спейсер и обратно.

В традиционной СС-схемотехнике одна ступень конвейера реализуется схемой, изображенной на рис. 6. Здесь индикаторные выходы комбинационной



**Рис. 5** Традиционное взаимодействие ступеней конвейера

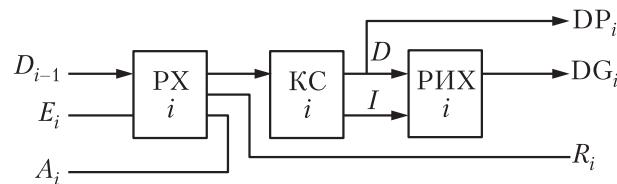


**Рис. 6** Структурная схема одной ступени традиционного конвейера

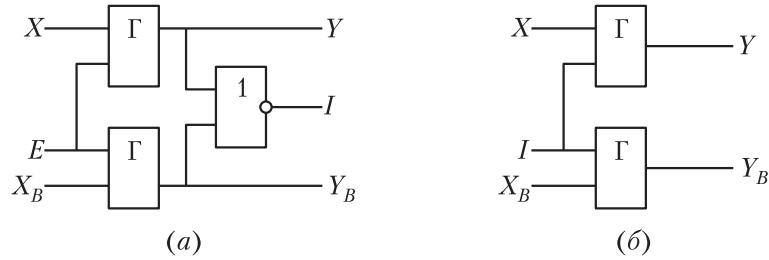
схемы (КС) и регистра хранения (РХ) объединяются в один индикаторный выходной сигнал всей ступени, который используется для организации взаимодействия с предыдущей и последующей ступенями конвейера. Однако в ряде случаев удается оптимизировать взаимодействие ступеней конвейера, повысив его быстродействие путем разделения индикаторных сигналов, управляющих предыдущей и последующей ступенями конвейера. В данном случае это оказалось возможным благодаря структуре каждой ступени конвейера, изображенной на рис. 7. Выходы  $DG_i$  индицируют поразрядные индикаторы и вместе с  $DP_i$  составляют общую совокупность парафазных информационных выходов ступени.

Каждая ступень конвейера, за исключением входного и выходного FIFO, включает три блока:

- (1) входной РХ;
- (2) КС, реализующую алгоритм обработки данных в текущей ступени;
- (3) выходной регистр индикации и хранения (РИХ).



**Рис. 7** Структурная схема одной ступени реализованного конвейера



**Рис. 8** Схема одного разряда регистров РХ (а) и РИХ (б)

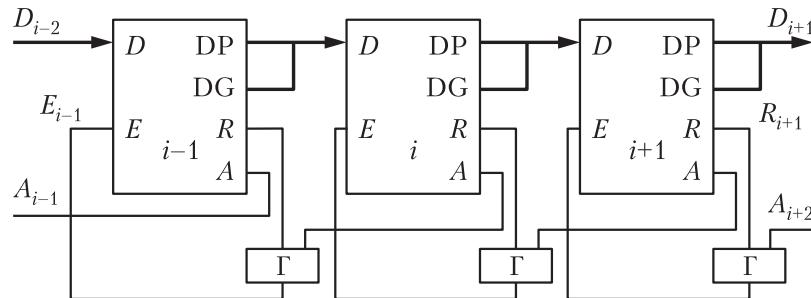
Регистр хранения формирует два индикаторных выхода:  $A_i$  и  $R_i$ , — причем сигнал  $A_i$ , как в данной реализации, может иметь спекулятивную основу, позволяющую ускорить запрос-ответное взаимодействие многоразрядных устройств в составе конвейера. Сигнал же  $R_i$  является полноценным индикатором РХ, обеспечивая строгую самосинхронность приема данных этой ступенью конвейера.

Реализация РХ (рис. 8, а) и РИХ (рис. 8, б) на основе Г-триггера позволяет оптимизировать по быстродействию индикаторную подсхему SIFMA. Здесь  $X$  и  $X_B$  — информационные парафазные входы;  $E$  — общий для всех разрядов РХ сигнал управления;  $I$  — индикаторный выход соответствующего разряда РХ или поразрядный индикатор комбинационной части данной ступени для РИХ;  $Y$  и  $Y_B$  — парафазные выходы с нулевым спейсером. Индикаторные поразрядные выходы РХ «сжимаются» в общий индикаторный выход  $R_i$  РХ с помощью стандартных схем индикации СС-схемотехники [10].

Регистры РХ и РИХ хранят как рабочую фазу, так и спейсер входных сигналов  $X$ ,  $X_B$ . Это избавляет от необходимости использовать дополнительно блок преобразования бифазного сигнала (выхода бистабильной ячейки триггера — разряда традиционного регистра хранения) в парафазный сигнал со спейсером, с которым работают комбинационные СС-схемы. Кроме того, РИХ индицирует своими выходами поразрядные индикаторы  $I$  схемы КС, а РХ — некоторый сигнал управления  $E$ , формируемый индикаторными выходами текущей и следующей ступеней конвейера. Тем самым упрощается и ускоряется формирование общего индикатора КС. Результирующая схема взаимодействия ступеней конвейера показана на рис. 9. Парафазные информационные выходы DP и DG объединяются в одну шину информационных данных.

Каждая ступень конвейера имеет спекулятивный (упрощенный на основе критического пути) и полный индикаторный выход. Спекулятивный индикаторный выход ступени  $A$  используется для формирования сигнала  $E$  управления предыдущей ступенью конвейера, а полный индикаторный выход  $R$  — для формирования сигнала управления текущей ступенью.

Такая реализация позволила достичь при типовых условиях средней производительности конвейера 1 Гфлопс и обеспечить среднее время выполнения операции не более 6 нс.



**Рис. 9** Оптимизированная схема взаимодействия ступеней конвейера SIFMA

Описанная организация запрос-ответного взаимодействия между ступенями конвейера была проанализирована программами анализа схемы на самосинхронность АСПЕКТ [13] и АСИАН [14], которые подтвердили принадлежность схемы конвейера к классу SI-устройств.

С учетом необходимости и достаточности индицирования в полном объеме только спейсерной фазы SIFMA индикаторная подсхема КС строится с использованием только комбинационных логических элементов, без гистерезисных триггеров, и реализует логическую функцию «И» (для единичного спейсера) либо «ИЛИ» (для нулевого спейсера). Это сокращает аппаратурные затраты и ускоряет формирование индикаторных сигналов, участвующих в запрос-ответном взаимодействии ступеней конвейера.

Разбиение SIFMA на ступени конвейера выполнялось исходя из принципа обеспечения максимального быстродействия SIFMA с учетом приемлемых аппаратурных затрат и сбалансированности конвейера. В целом оно соответствует структурной схеме SIFMA на рис. 3. В варианте с синхронным окружением входное и выходное FIFO являются дополнительными ступенями конвейера.

Конвейер SIFMA обеспечивает производительность на уровне 1 Гфлопс для среднестатистической комбинации входных операндов при типовом напряжении питания 1 В и температуре окружающей среды 25 °C. Среднее время выполнения операции — не более 6 нс. При обработке наихудшей с точки зрения времени выполнения комбинации входных operandов производительность может снижаться, но в среднем она будет не хуже 1 Гфлопс.

#### 4 Характеристики SIFMA

Характеристики любого устройства во многом определяются базисом его реализации. В настоящее время известны два основных базиса реализации SI-схем:

- (1) избыточная NCL (NULL convention logic) логика [15], обеспечивающая независимость не только от задержек элементов, но и от задержек в цепях межсоединений;

## Характеристики вариантов SIFMA

Интерфейс SIFMA	Сложность реализации, транзисторы	Площадь, мм	Производительность, Гфлопс	Время выполнения операции, нс	Энергопотребление, мДж / ГГц
A	639 000	0,78	1,0	5,95	970
C	724 000	0,96	1,0	6,90	1140

- (2) неизбыточная КМОП-логика [16], обеспечивающая независимость от задержек элементов, а в пределах эквихронной зоны [10] — и от задержек в цепях межсоединений.

После сравнения преимуществ и недостатков обоих схемотехнических базисов [16] предпочтение было отдано неизбыточной КМОП-логике, так как она обеспечивает меньшие аппаратурные затраты (в 4,49 раза при реализации двоичного счетчика, в 1,13 раза при реализации умножителя  $4 \times 4$ , до 2 раз при реализации более простых логических схем), большую производительность и меньшее энергопотребление по сравнению с NCL-схемами.

При проектировании SIFMA использовалась библиотека элементов, разработанная для стандартной КМОП-технологии с проектными нормами 65 нм. Двадцать семь элементов, вошедших в состав библиотеки, являются подмножеством элементов из библиотеки для проектирования СС-схем [17]. Они были апробированы при разработке и изготовлении по КМОП-технологии 65-нанометровыми проектными нормами SI-вычислителя [9].

Характеристики вариантов SIFMA для асинхронного (A) и синхронного (C) окружения, выполненных по 65-нанометровой КМОП-технологии с шестью слоями металлизации, приведены в таблице. Они являются оценочными. Они получены путем моделирования схемы SIFMA с учетом паразитных параметров, извлеченных из топологии блока. Быстродействие определялось для типовых условий эксплуатации, так как производительность SI-схем всегда соответствует текущим условиям эксплуатации, а сами SI-схемы не требуют учета наихудшего случая.

В настоящее время вариант SIFMA для синхронного окружения находится на этапе завершения топологического проектирования и готовится к запуску в составе тестовой БИС (большой интегральной схемы).

## 5 Заключение

Представленные варианты устройства, выполняющего операцию FMA в соответствии со стандартом IEEE 754, относятся к классу устройств, поведение которых не зависит от задержек элементов (SI-устройств). Они позволяют получить сумму и разность третьего операнда и произведения двух первых операндов для одной тройки чисел двойной точности или для двух упакованных троек чисел одинарной точности.

Вариант с синхронным окружением отличается от варианта с асинхронным окружением наличием входного и выходного FIFO емкостью в 4 слова данных каждое.

Средняя производительность обоих вариантов устройства при типовых условиях равна 1,0 Гфлопс.

Энергопотребление SIFMA составляет 970 и 1140 мДж / ГГц для вариантов с асинхронным и синхронным окружением соответственно.

В настоящее время вариант SIFMA для синхронного окружения готовится к запуску в составе тестовой БИС.

## Литература

1. *Chen C., Chen L.-A., Cheng J.-R.* Architectural design of a fast floating-point multiplication-add fused unit using signed-digit addition // IEE Proc. Comput. Dig. Tec., 2002. Vol. 149. No. 4. P. 113–120.
2. *Seidel P.* Multiple path IEEE floating-point fused multiply-add // 46th IEEE Midwest Symposium (International) on Circuits and Systems Proceedings. — Cairo, Egypt, 2003. P. 1359–1362.
3. *Lang T., Bruguera J.* Floating-point fused multiply-add with reduced latency // IEEE Trans. Comput., 2004. Vol. 53. No. 8. P. 42–51.
4. *Libo H., Li S., Kui D., Zhiying W.* A new architecture for multiple-precision floating-point multiply-add fused unit design // 18th IEEE Symposium on Computer Arithmetic Proceedings. — Montpellier, 2007. P. 69–76.
5. *Quinnell E. C., Swartzlander E., Lemonds C.* Floating-point fused multiply-add architectures // 41th Asilomar Conference on Signals, Systems and Computers Proceedings. — Pacific Grove, États-Unis, 2007. P. 42–51.
6. *Quinnell E. C., Swartzlander E., Lemonds C.* Three-path fused multiply adder circuit. United States Patent Application Publication No. 0256150. 2008.
7. *Walaa Abd El Aziz I.* Binary floating point fused multiply add unit. Master Thesis. — Giza, Egypt, 2012. 88 p.
8. IEEE Standard for floating-point arithmetic 754-2008. — IEEE, 2008. doi: 10.1109/IEEEESTD.2008.4610935.
9. Степченков Ю. А., Дьяченко Ю. Г., Рождественский Ю. В., Морозов Н. В., Степченков Д. Ю. Разработка вычислителя, не зависящего от задержек элементов // Системы и средства информатики, 2010. Вып. 20. № 1. С. 5–23.
10. Автоматное управление асинхронными процессами в ЭВМ и дискретных системах / Под ред. В. И. Варшавского. — М.: Наука, 1986. 400 с.
11. *Makino H., Nakase Y., Suzuki H., Morinaka H., Shinohara H., Mashiko K.* An 8.8 ns 54 × 54 bit multiplier with high speed redundant binary architecture // IEEE J. Solid-St. Circ., 1996. Vol. 31. No. 6. P. 773–783.
12. *Hensley J., Lastra A., Singh M.* A scalable counterflow-pipelined asynchronous radix-4 booth multiplier // Symposium (International) on Asynchronous Circuits and Systems Proceedings, 2005. P. 128–137.

13. Рождественский Ю. В., Морозов Н. В., Рождественскене А. В. АСПЕКТ: Подсистема событийного анализа самосинхронных схем // Проблемы разработки перспективных микро- и наноэлектронных систем: IV Всеросс. науч.-технич. конф. (МЭС-2010). — М.: ИППМ РАН, 2010. С. 26–31.
14. Рождественский Ю. В., Морозов Н. В., Степченков Ю. А., Рождественскене А. В. Универсальная подсистема анализа самосинхронных схем // Системы и средства информатики, 2006. Вып. 16. № 2. С. 463–475.
15. Fant K. M. Logically determined design: Clockless system design with NULL convention logic. — New York: J. Wiley & Sons, 2005. 292 р.
16. Соколов И. А., Степченков Ю. А., Бобков С. Г., Захаров В. Н., Дьяченко Ю. Г., Рождественский Ю. В., Сурков А. В. Базис реализации супер-ЭВМ эксафлопсного класса // Информатика и её применения, 2014. Т. 8. Вып. 1. С. 45–70.
17. Степченков Ю. А., Денисов А. Н., Дьяченко Ю. Г. и др. Библиотека элементов для проектирования самосинхронных полуузаказных БМК микросхем серий 5503/5507 и 5508/5509. — М.: ИПИ РАН, 2013. 391 с.

*Поступила в редакцию 20.08.14*

---

## SELF-TIMED FUSED MULTIPLY-ADD UNIT: PRACTICAL IMPLEMENTATION

***Y. Stepchenkov<sup>1</sup>, Y. Diachenko<sup>1</sup>, Y. Rogdestvenski<sup>1</sup>, N. Morozov<sup>1</sup>,  
D. Stepchenkov<sup>1</sup>, A. Rogdestvenskene<sup>1</sup>, and A. Surkov<sup>2</sup>***

<sup>1</sup>Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333,  
44-2 Vavilova Str., Russian Federation

<sup>2</sup>Scientific Research Institute for System Studies, Russian Academy of Sciences,  
36 bld. 1 Nakhimovsky Prospekt, Moscow 117218, Russian Federation

**Abstract:** Paper presents the results of designing Speed-Independent Fused Multiply-Add (SIFMA) variants compliant with the IEEE 754 Standard. SIFMA performs either one double precision FMA operation or one or two single precision operations simultaneously under three operands. SIFMA was designed for the standard 65-nanometer CMOS (complementary metall-oxide-semiconductor) technology. It uses both a standard cell library and a self-timed cell library designed in IPI RAS. One SIFMA variant operates with a synchronous environment, while another works with an asynchronous environment. Both variants provide an average performance up to 1 GFlops for 1-volt supply and environment temperature of 25 °C. At these conditions, energy consumption does not exceed 970 mJ/GHz.

**Keywords:** self-timed circuit; ternary coding; multiplier; adder; subtracter; pipeline; indication

**DOI:** 10.14357/08696527140305

## Acknowledgments

This project was financially supported by the Russian Foundation for Basic Research (project 13-07-12068 ofi\_m).

## References

1. Chen, C., L.-A. Chen, and J.-R. Cheng. 2002. Architectural design of a fast floating-point multiplication-add fused unit using signed-digit addition. *IEE Proc. Comput. Dig. Tec.* 149(4):113–120.
2. Seidel, P. 2003. Multiple path IEEE floating-point fused multiply-add. *46th IEEE Midwest Symposium (International) on Circuits and Systems Proceedings*. Cairo, Egypt. 1359–1362.
3. Lang, T., and J. Bruguera. 2004. Floating-point fused multiply-add with reduced latency. *IEEE Trans. Comput.* 53(8):42–51.
4. Libo, H., S. Li, D. Kui, and W. Zhiying. 2007. A new architecture for multiple-precision floating-point multiply-add fused unit design. *18th IEEE Symposium on Computer Arithmetic Proceedings*. Montpellier. 69–76.
5. Quinnell, E. C., E. Swartzlander, and C. Lemonds. 2007. Floating-point fused multiply-add architectures. *41th Asilomar Conference on Signals, Systems and Computers Proceedings*. Pacific Grove, États-Unis. 42–51.
6. Quinnell, E. C., E. Swartzlander, and C. Lemonds. 2008. Three-path fused multiply adder circuit. United States Patent Application Publication No. 0256150.
7. Walaa Abd El Aziz, I. 2012. Binary floating point fused multiply add unit. Master Thesis. Giza, Egypt. 88 p.
8. IEEE Computer Society. 2008. IEEE standard for floating-point arithmetic 754-2008. doi: 10.1109/IEEEESTD.2008.4610935.
9. Stepchenkov, Y. A., Y. G. Diachenko, Y. V. Rogdestvenski, N. V. Morozov, and D. Y. Stepchenkov. 2010. Razrabotka vychislitelya, nezavisimogo ot zaderzhek elementov [Designing of the delay independent computing device]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 20(1):237–254.
10. Varshavsky, V. I., ed. 1990. *Self-timed control of concurrent processes: The design of aperiodic logical circuits in computers and discrete systems*. Mathematics and its applications ser. Dordrecht, The Netherlands: Kluwer Academic Publs. 418 p.
11. Makino, H., Y. Nakase, H. Suzuki, H. Morinaka, H. Shinohara, and K. Mashiko. 1996. An 8.8 ns 54 × 54 bit multiplier with high speed redundant binary architecture. *IEEE J. Solid-St. Circ.* 31(6):773–783.
12. Hensley, J., A. Lastra, and M. Singh. 2005. A scalable counterflow-pipelined asynchronous radix-4 booth multiplier. *Symposium (International) on Asynchronous Circuits and Systems Proceedings*. 128–137.
13. Rogdestvenski, Y. V., N. V. Morozov, and A. Rogdestvenskene. 2010. ASPEKT: Podistema sobitynogo analiza samosinkronnykh skhem [ASPECT: A suite of self-timed event-driven analysis]. *Tr. Mezhdunar. konf. “Problemy Razrabotki Perspektivnykh mikro- i nanoelektronnykh sistem”* [Problems of the Perspective Micro- and Nanoelectronic Systems Development Proceedings]. Moscow. 26–31.

14. Rogdestvenski, Y. V., N. V. Morozov, Y. A. Stepchenkov, and A. V. Rogdestvenskene. 2006. Universal'naya podsistema analiza samosinkhronnykh skhem [The universal suite for self-timed circuit analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 16(2):463–475.
15. Fant, K. M. 2005. *Logically determined design: Clockless system design with NULL convention logic*. New York: J. Wiley & Sons. 292 p.
16. Sokolov, I. A., Y. A. Stepchenkov, S. G. Bobkov, V. N. Zakharov, Y. G. Diachenko, Y. V. Rogdestvenski, and A. V. Surkov. 2014. Bazis realizatsii super-EVM eksaflopsnogo klassa [Implementation basis of ExaFlops class supercomputer]. *Informatika i ee Primeneniya — Inform. Appl.* 8(1):45–70.
17. Stepchenkov, Y. A., A. N. Denisov, Y. G. Diachenko, et al. 2013. *Biblioteka elementov dlya proektirovaniya samosinkhronnykh poluzakaznykh BMK seriy 5503/5507 i 5508/5509* [Cell library for designing self-timed 5503/5507 and 5508/5509 gate-arrays]. Moscow: IPI RAN. 391 p.

*Received August 20, 2014*

## Contributors

**Stepchenkov Yuri A.** (b. 1951) — Candidate of Science (PhD) in technology, Head of Department, Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333, Russian Federation; YStepchenkov@ipiran.ru

**Diachenko Yuri G.** (b. 1958) — Candidate of Science (PhD) in technology, senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333, Russian Federation; diaura@mail.ru

**Rogdestvenski Yuri V.** (b. 1952) — Candidate of Science (PhD) in technology, Head of Laboratory, Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333, Russian Federation; YRogdest@ipiran.ru

**Morozov Nikolay V.** (b. 1956) — senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333, Russian Federation; NMorozov@ipiran.ru

**Stepchenkov Dmitri Y.** (b. 1973) — senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333, Russian Federation; Stepchenkov@mail.ru

**Rogdestvenskene Asta V.** (b. 1964) — senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, Moscow 119333, Russian Federation; astarozd@yahoo.com

**Surkov Alexei V.** (b. 1978) — senior scientist, Scientific Research Institute for System Studies, Russian Academy of Sciences, Moscow 117218, Russian Federation; surkov@cs.niisi.ras.ru

## СОВРЕМЕННЫЕ ТЕНДЕНЦИИ В РАЗВИТИИ АРХИТЕКТУР ИНТЕГРИРОВАННЫХ СЕТЕВЫХ ПРОЦЕССОРОВ

*В. Б. Егоров<sup>1</sup>*

**Аннотация:** В статье отмечена возрастающая роль программируемых интегрированных сетевых процессоров (ИСП) в создании современных пакетных коммутаторов, маршрутизаторов и различных устройств сетевой инфраструктуры, дан анализ сегодняшних тенденций в развитии архитектур ИСП, выделены различные уровни программируемости и важнейшие компоненты макроархитектуры ИСП, выявлена взаимосвязь макроархитектуры с возможностью организации в ИСП внутренних виртуальных конвейеров для детерминистической обработки пакетов с высокой пропускной способностью, показана ключевая роль диспетчера внутренних очередей в достижении эффективного функционирования внутри ИСП виртуальных конвейеров, отмечены преимущества высокоскоростных последовательных интерфейсов, характеризующихся организацией взаимодействия между абонентами на равноправной основе в качестве средства подключения множества ИСП к общей внешней коммутационной структуре.

**Ключевые слова:** интегрированный сетевой процессор; многоядерный процессор; макроархитектура; виртуальный конвейер; диспетчер очередей; высокоскоростные последовательные интерфейсы

**DOI:** 10.14357/08696527140306

### 1 Введение

По мере расширения пакетных сетей до регионального и глобального масштаба усложняется их топология, коммутируемым пакетам приходится последовательно проходить все большее число сетевых узлов, что порождает стремление сократить время продвижения пакетов внутри отдельного узла. В прошлом веке оно выразилось тенденцией к замене традиционных программных методов коммутации (soft switching) аппаратными (hard forwarding). Наиболее ярко это стремление проявилась в переходе от технологии Frame Relay к Asynchronous Transfer Mode (ATM). Аппаратная коммутация в совокупности с достижениями микроэлектроники позволила реализовать пакетный коммутатор на одном кристалле кремния [1, 2]. Однако век ATM оказался недолгим. Очень скоро, уже в конце прошлого столетия, продолжающаяся глобализация сетей и стремительный рост в них числа абонентов стали выдвигать на передний план требования обеспечения качества обслуживания и информационной безопасности в сетях —

---

<sup>1</sup>Институт проблем информатики Российской академии наук, VEgorov@ipiran.ru

свойств, реализация которых основывается на комплексах сложных, постоянно совершенствующихся протоколов. Необходимость поддержки в современных сетях множества разнообразных и постоянно обновляемых протоколов сдвинула приоритеты в реализации узлов пакетных сетей в сторону программируемых устройств. При этом их производители не хотели отказываться от преимуществ однокристальной реализации, что заставило микроэлектронные фирмы искать принципиально новое решение, которое было найдено в высокопроизводительных ИСП, реализуемых как интеллектуальные системы на кристалле. Интеллект таких систем обеспечивается программируемыми многоядерными процессорами, а высокая производительность достигается за счет большого числа универсальных процессорных ядер, число которых может исчисляться десятками [3–5], и дополнительных специализированных функциональных блоков, как ограниченно программируемых, так и чисто аппаратных [6].

## 2 Макроархитектура интегрированного сетевого процессора

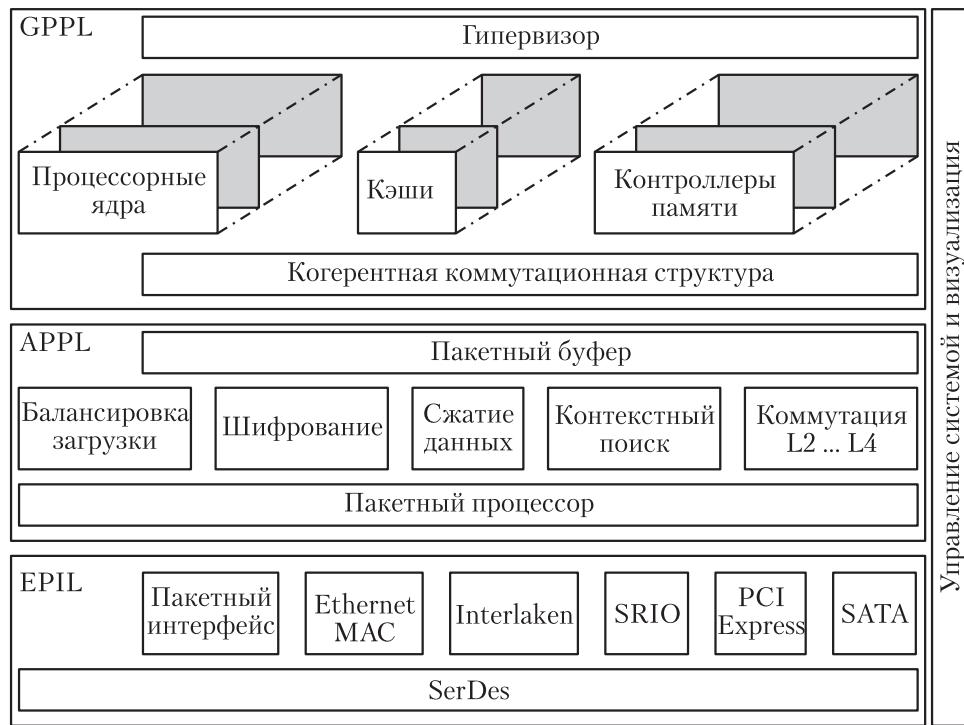
Хорошой иллюстрацией современных тенденций в развитии архитектур ИСП может служить анонсированная в 2012 г. компанией Freescale Semiconductor макроархитектура Layerscape [7, 8], она показана на рис. 1.

Главное достоинство этой макроархитектуры заключается в ее существенной независимости от системы инструкций ISA (instruction set architecture) процессорных ядер и функциональных особенностей специализированных блоков и аппаратных ускорителей. С точки зрения программируемости в макроархитектуре Layerscape выделяются три уровня:

- (1) универсальной обработки GPPL (general-purpose processing layer);
- (2) ускоренной обработки пакетов APPL (accelerated packet processing layer);
- (3) «курьерского» ввода–вывода пакетов EPIL (express packet input/output (I/O) layer).

Свободно программируемый уровень GPPL включает произвольное множество процессорных ядер с любой ISA, подсистему памяти из нескольких кэшей двух–трех уровней и контроллеров внешней оперативной памяти, а также когерентную коммутационную структуру, обеспечивающую бесконфликтный трафик между процессорными ядрами и подсистемой памяти с поддержкой когерентности кэшей. Номинальная производительность конкретной конфигурации уровня GPPL определяется числом ядер и контроллеров памяти при соответствующей пропускной способности коммутационной структуры. Фактическая производительность уровня в общем случае сильно зависит от возможностей гипервизора и эффективности поддержки имеющимися операционными системами (ОС) симметричной многопроцессорности (СМП).

Уровень APPL, основным элементом которого является пакетный процессор, может выполнять вспомогательные задачи, привлекая по мере надобности

**Рис. 1** Макроархитектура Layerscape

специализированные функциональные блоки: шифрования, компрессии и де-компрессии данных, ускорения табличного и контекстного поиска, балансировки загрузки и др. Уровень APPL не только снимает часть нагрузки с уровня GPPL, но также способен самостоятельно осуществлять манипуляции с пакетами по типовым алгоритмам, например коммутацию на канальном и сетевом уровнях. Компоненты APPL могут программироваться пользователем, в том числе структурным программированием на основе языка Си, но с известными ограничениями, накладываемыми специализацией аппаратуры. Вероятно, ОС для уровня APPL должны так или иначе поддерживать асимметричную многопроцессорность (АМП).

Компоненты уровня EPIL, отвечающего за ввод-вывод, скорее программируются, чем конфигурируются. Пакетный интерфейс выполняет классификацию входящих кадров или пакетов и принимает решение об их направлении по тому или иному маршруту обработки внутри ИСП в соответствии с предварительными настройками. Помимо него в аппаратуру уровня EPIL входят контроллеры основных сетевых интерфейсов, как правило, высокоскоростного Ethernet, контроллеры служебных коммуникационных интерфейсов

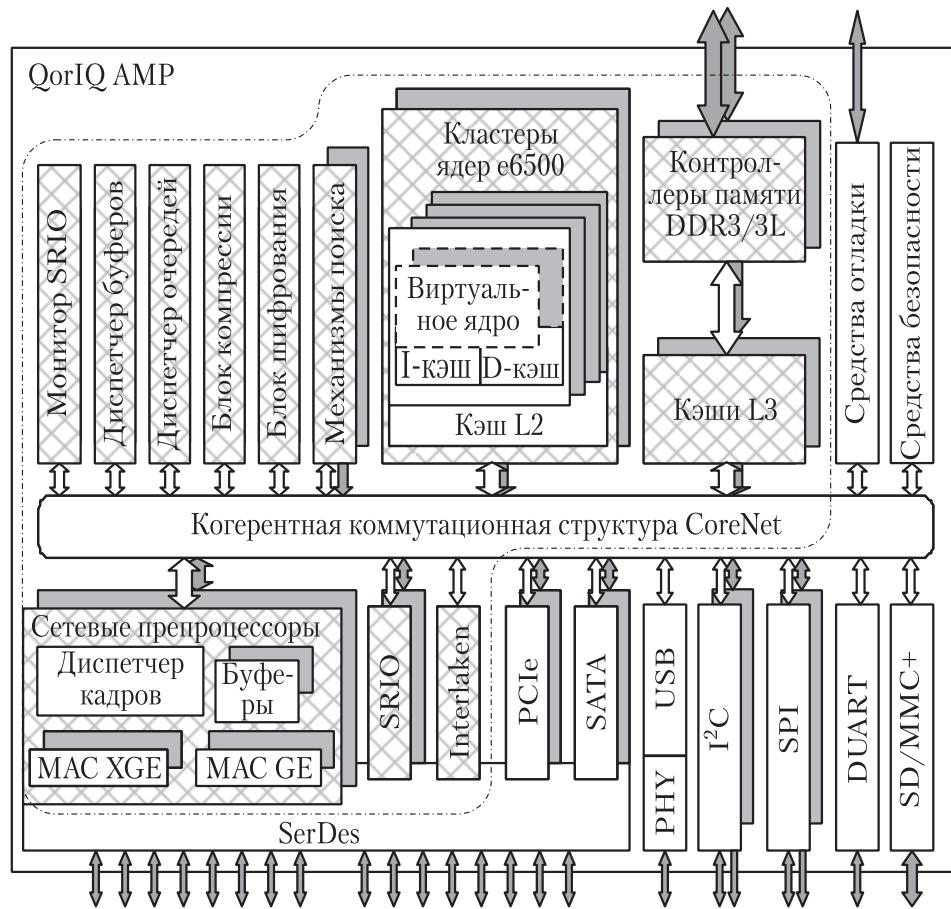
Interlaken и SRIO (Serial Rapid I/O), а также общесистемных интерфейсов PCI (Peripheral Component Interconnect) Express (PCIe), SATA (Serial Advanced Technology Attachment) и USB (Universal Serial Bus). Поскольку макроархитектура ориентирована на интегральную реализацию, в целях экономии выводов корпуса ИСП предпочтение в ней отдается высокоскоростным последовательным интерфейсам, обслуживаемым универсальным преобразователем SerDes (Serializer / Deserializer) параллельных кодов в последовательные и обратно.

Концептуально прикладное программирование в макроархитектуре Layerscape обретает новое качество вследствие независимости от ISA центрального процессора и унификации модели программирования. Любое приложение предполагается равно исполнимым на разных моделях ИСП с различными ISA и конфигурациями процессора: от простейших одноядерных с жесткими ограничениями по потребляемой мощности в мобильных приложениях до высокопроизводительных многоядерных в сетевых серверах. Макроархитектура поддерживает современные методы программирования с открытыми стандартами (open standard software programming) и учитывает последние ключевые новшества в промышленном программном обеспечении (ПО), включая программно определяемые сети SDN (Software Defined Networks).

### 3 Виртуальный конвейер

К сожалению, на сегодняшний день макроархитектура Layerscape — скорее абстрактная концепция, чем руководство к действию. Воплощение она нашла лишь в небольшом семействе ИСП QorIQ LS [8], причем в сильно урезанном виде: с минимальным числом процессорных ядер на уровне GPPL и вырожденными уровнями APPL и EPIL. Но потенциал подобных макроархитектур раскрылся уже в предшественнице Layerscape — макроархитектуре ускорения путей данных DPAA (Data Paths Acceleration Architecture) [9], реализованной, в частности, в обширном семействе ИСП QorIQ AMP (Advanced MultiProcessing) [10]. Обобщенная структура ИСП этого семейства показана на рис. 2.

В семействе ИСП QorIQ AMP уровень GPPL макроархитектуры Layerscape представлен несколькими кластерами из четырех процессорных ядер e6500 с ISA PowerPC, причем сдвоенный конвейер каждого RISC (Restricted Instruction Set Computing) ядра виртуально эмулирует два процессора, вследствие чего эффективное число ядер в кластере оказывается равным восьми. Кластер снабжен общим для всех ядер кэшем второго уровня. К уровню APPL относятся диспетчеры очередей и буферов, механизмы табличного и контекстного поиска, блоки шифрования, компрессии и декомпрессии данных. Уровень EPIL включает множество разнообразных стандартных интерфейсов и специализированный сетевой процессор с несколькими MAC-контроллерами Ethernet для скоростей 1 Гбит/с (GE — Gigabit Ethernet) и 10 Гбит/с (XGE), а также диспетчер кадров, выполняющий функции пакетного интерфейса в макроархитектуре Layerscape.

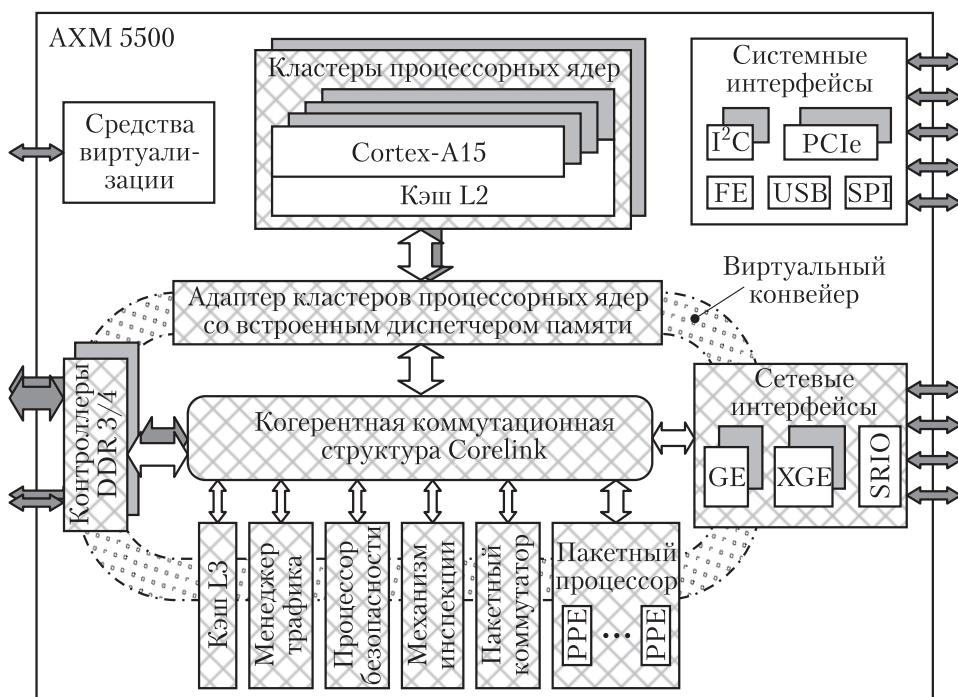


**Рис. 2** Обобщенная структура ИСП семейства QorIQ AMP (DDR — Double Data Rate; I<sup>2</sup>C — InterIntegrated Circuit; SPI — Serial Peripheral Interface; DUART — Dual Universal Asynchronous Receiver/Transmitter; SD — Secure Digital memory card; MMC — MultiMedia Card)

Основное достоинство макроархитектуры DPAA, как и Layerscape, заключается в уменьшении нагрузки на прикладные программы, т. е. на уровень GPPL Layerscape, за счет «автоматического» подключения к процессу обработки кадров и пакетов аппаратных блоков уровня APPL. Такое подключение осуществляется тремя встроенными в ИСП диспетчерами: кадров, буферов и очередей. Диспетчер кадров непосредственно после приема и буферизации классифицирует очередной кадр и «прописывает» ему некий маршрут обработки внутри ИСП, который может задействовать те или иные блоки уровней GPPL и APPL (эти

блоки на рис. 2 отмечены «сетчатой» штриховкой и обведены штрихпунктиром). Прохождение пакета по назначенному маршруту контролируется диспетчером очередей. Одни маршруты, требующие специфической обработки прикладными программами, включают универсальные процессорные ядра; другие, свойственные типовым функциям обработки пакетов и целиком реализуемые на уровне APPL, например коммутация кадров, могут вообще не затрагивать процессор. Благодаря автоматическому глубокому буферированию пакетов и наличию программируемого диспетчера очередей появляется возможность применять внутри ИСП общепринятые механизмы обеспечения качества обслуживания и балансировки загрузки компонентов DPAA.

Фактически внутри ИСП организуется некий виртуальный конвейер со специализированными рабочими местами, даже несколько таких конвейеров соответственно прописанным маршрутам обработки. Методологический взгляд на такого рода пакетную обработку внутри ИСП как на виртуальный конвейер предложила компания LSI в своих новейших подсемействах AXE 4500 и AXM 5500 семейства ИСП Axxia [11]. Структура ИСП AXE 5500 качественно не отличается от



**Рис. 3** Виртуальный конвейер в ИСП подсемейства Axxia AXM 5500 (FE — fast Ethernet)

структуры ИСП QorIQ AMP за исключением ISA процессора, а количественные различия могут касаться числа кластеров ядер или механизмов обработки пакетов PPE (Packet Processing Engines) в пакетном процессоре. В виртуальном конвейере задействуются те же самые блоки, что и в макроархитектуре DPAA. На рис. 3, показывающем организацию виртуального конвейера ИСП AXE 5500, этот факт подчеркивается той же их штриховкой этих блоков, как и на рис. 2. К сожалению, LSI не раскрывает практические механизмы реализации своего виртуального конвейера, но вряд ли они существенно отличаются от механизмов, использованных компанией Freescale в DPAA.

#### **4 Диспетчер очередей**

В концепции макроархитектуры DPAA (и, вероятно, в реализации виртуального конвейера) важная роль отводится диспетчеру очередей. Эта роль определяется двумя взаимосвязанными причинами: необходимостью собственно диспетчеризации очередей пакетов на обслуживание к различным макроархитектурным компонентам ИСП и обеспечением качества этого обслуживания в процессе диспетчеризации. Польза специального выделенного диспетчера ярко проявляется, в частности, в управлении многовходовыми очередями, типичным примером которых могут служить очереди пакетов к выходным сетевым портам. Количество таких очередей в ИСП равно произведению числа портов на количество установленных уровней приоритета исходящих пакетов.

Пусть в некоторую очередь  $Q$  в любой момент может быть вставлен пакет из нескольких источников. Например, плоскость данных может поставлять коммутируемые пакеты из входных портов либо зашифрованные пакеты из блока шифрования, а плоскость управления — порожденные внутри ИСП пакеты протокола маршрутизации. Таким образом, источниками пакетов в одну и ту же многовходовую очередь могут служить различные плоскости сервисов и физически разные блоки ИСП, процессорные ядра и аппаратные ускорители. Пример многовходовой очереди  $Q$  приведен на рис. 4, где предполагается, что пакеты, предварительно буферированные в памяти входными портами ИСП с участием диспетчера буферов, представлены в очереди своими описателями, которые могут содержать адрес буфера, длину пакета и другие параметры.

Поскольку в общем случае все поставщики пакетов в очередь  $Q$  независимы, принципиально существует вероятность попытки внести элемент в очередь одновременно двумя и более поставщиками, что приводит к конфликту между ними. Этот конфликт не может быть разрешен на приоритетной основе, так как все вносимые в одну очередь пакеты равноприоритетны. Разрешение конфликта традиционными методами управления доступом к разделяемому ресурсу, каковым является очередь  $Q$ , например механизмом семафоров, не решает проблему первенства и ведет в лучшем случае к снижению пропускной способности из-за простоя на семафорах, а в худшем — к неразрешимым блокировкам (deadlocks), полностью парализующим работу устройства.

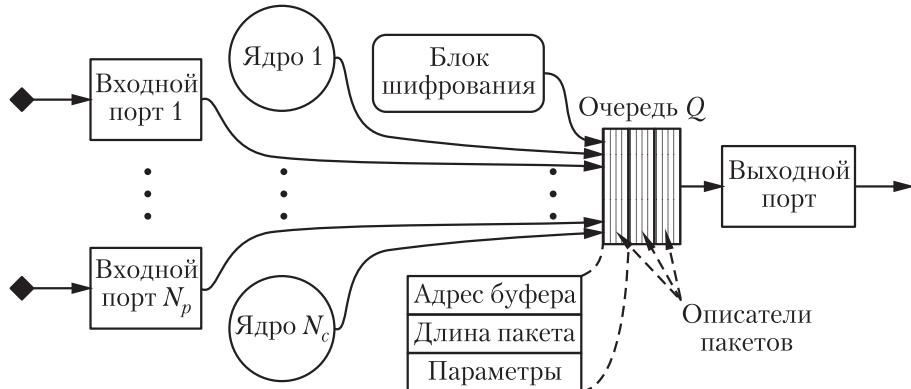


Рис. 4 Пример многовходовой очереди

Типичный механизм возникновения неразрешимой блокировки на семафоре иллюстрируется следующим примером. Пусть некая высокоприоритетная программа  $P_H$  прерывает некоторую низкоприоритетную программу  $P_L$ , когда последняя находится в процессе внесения элемента в очередь  $Q$  и на момент прерывания очередной элемент внесен программой  $P_L$  лишь частично. Далее, пусть программе  $P_H$  также требуется внести элемент в ту же очередь  $Q$ . Очевидно, она не сможет этого сделать, не нарушив структуру очереди, поэтому очередь  $Q$  как разделенный ресурс должна быть недоступна программе  $P_H$ , например, из-за блокирующего доступа к очереди  $Q$  семафора, установленного программой  $P_L$ . Но программа  $P_L$  не в состоянии завершить процесс внесения своего элемента в очередь  $Q$  и открыть семафор, поскольку в свою очередь заблокирована прервавшей ее программой  $P_H$ .

Проблема решается разделением многовходовой очереди  $Q$  с  $K$  входами на  $K$  одновходовых очередей  $Q_k$  ( $k = 1, \dots, K$ ) и передачей функции объединения всех очередей  $Q_k$  в одну очередь  $Q$  диспетчеру очередей, как показано на рис. 5.

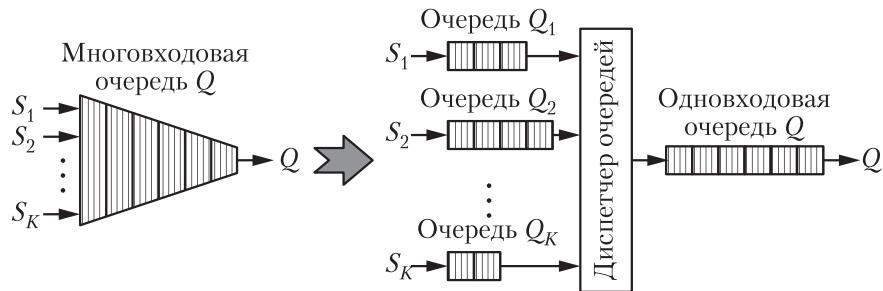


Рис. 5 Объединение нескольких очередей в одну диспетчером очередей

Помимо решения проблемы неразрешимых блокировок на входах многовходовых очередей специализированный диспетчер может обеспечить выполнение требований качества обслуживания дифференцированно для различных потоков пакетов внутри ИСП. При этом могут реализовываться такие общепринятые алгоритмы, как полисинг на входах всех очередей со случайным ранним отбрасыванием RED (Random Early Discard) избыточных пакетов, взвешенный круговой выбор WRR (Weighted Round Robin) пакетов на выходах очередей  $Q_k$  и шейпинг потока на выходе очереди  $Q$ .

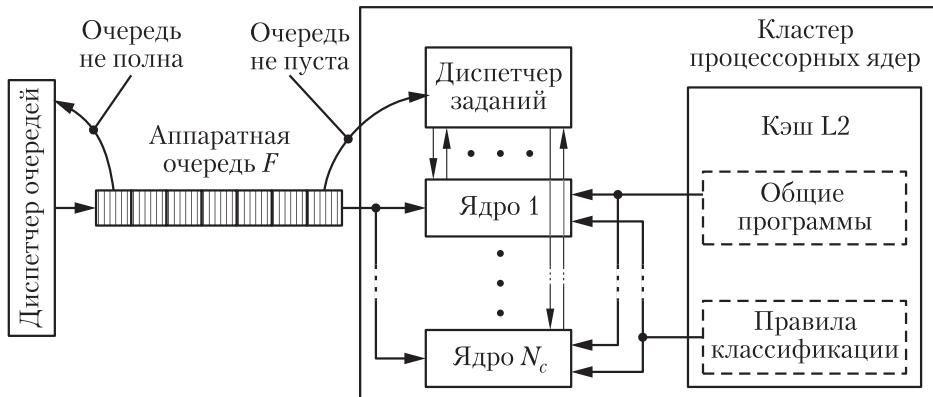
## 5 Специализация кластеров процессорных ядер

Как видно из рис. 2 и 3, наращивание числа процессорных ядер в реальных ИСП сопровождается их кластеризацией, которая дает возможность сократить количество кэшей второго уровня и, как следствие, абонентов когерентной коммутационной структуры. Но, с другой стороны, разделяемый всеми ядрами кластера кэш L2 будет эффективно использоваться только при общем для всех ядер ПО и существенно общих используемых данных, т. е. в конечном счете при функциональной специализации кластеров. Такая специализация на кластерном уровне не противоречит ни макроархитектурам DPAA и Layerscape, ни концепции виртуального конвейера. С учетом решаемых ИСП задач возможны два подхода к организации вычислительного процесса внутри функционально специализированного кластера процессорных ядер.

Первый подход, успешно применяемый в универсальных многоядерных процессорах, основывается на традиционной СМП при мощной общей для всех ядер многозадачной и / или многопоточной ОС (ОС реального времени (РВ)). Очевидное достоинство данного подхода заключается в его апробированности, а существенный недостаток — в больших и трудно предсказуемых накладных расходах ОС на переключение задач и передачу их другим ядрам, заметно снижающих общую производительность процессора и лишающих весь вычислительный процесс детерминизма.

Второй подход, учитывающий специфику ИСП, предполагает независимую работу ядер кластера. Все ядра решают в целом одну и ту же задачу одним и тем же прикладным ПО с общими базами данных, но каждое ядро решает ее от начала до конца со своим комплектом текущих переменных. При этом диспетчеризация заданий между ядрами кластера может быть аппаратной, что делает вообще необязательным наличие в кластере какой-либо ОС, а отсутствие пересечений по текущим переменным исключает необходимость в поддержании когерентности кэшей L2. Хорошим примером применения такого подхода может стать показанный на рис. 6 кластер классификации входящих пакетов.

Аппаратная очередь  $F$  логически представляет собой голову очереди на классификацию пакетов, создаваемой диспетчером очередей, а физически — тривиальное FIFO (First In, First Out) с индикаторами его непустоты и неполноты. Структура элемента очереди  $F$ , т. е. описателя пакета, в основном



**Рис. 6** Пример организации кластера классификации пакетов

та же, что и на рис. 5. Пока очередь не полна, диспетчер очередей имеет возможность добавлять в нее новые элементы. Если очередь не пуста, то диспетчер заданий кластера, роль которого в простейшем случае может исполнять контроллер прерываний, дает команду (прерывание) свободному процессорному ядру на извлечение из очереди  $F$  описателя, после чего ядро может приступить к классификации буферированного пакета, на который ссылается этот описатель. Результатом классификации, отправляемым обратно диспетчеру очередей, будет тот же самый описатель с добавленным в него маршрутом обработки пакета внутри ИСП или, что фактически то же самое, конфигурацией виртуального конвейера.

Маршрут может быть задан явно, подобно стеку меток MPLS (MultiProtocol Label Switching), списком очередей, через которые последовательно должен пройти данный пакет. Но при этом размеры описателей пакетов окажутся переменными, что создаст неудобства с их размещением в очередях, особенно аппаратных (типа очереди  $F$  на рис. 6). Для фиксации размера и формата описателя маршрут можно задавать коротким номером, а соотнесенную с ним последовательность очередей хранить в небольшого объема маршрутной таблице внутри диспетчера очередей.

По мере развития реальных ИСП некоторые свободно программируемые специализированные процессорные кластеры могут заменяться функционально эквивалентными, но более производительными аппаратными ускорителями. Существенно, что такие замены никак не отразятся на работе диспетчера очередей и функционировании виртуальных конвейеров, а при грамотно организованных интерфейсах прикладного программирования API (Application Programming Interfaces) эти замены не повлияют и на прикладное ПО.

Серьезной проблемой кластерной организации ИСП может оказаться взаимодействие между собой программ, выполняемых в разных процессорных класте-

рах. Эффективность такого взаимодействия на общем поле оперативной памяти, доступной всем ядрам через когерентную коммутационную структуру, может оказаться чрезвычайно низкой из-за больших задержек доступа и многочисленных конфликтов, вплоть до неразрешимых, на разделяемых ресурсах памяти. Поэтому более эффективной, если не единственной реализуемой в ИСП, может оказаться организация взаимодействия между собой программ, исполняемых в разных кластерах, на сетевых принципах — обменом сообщениями, задействуя для этого диспетчер очередей и типовые механизмы обеспечения качества обслуживания.

## **6 Интерфейсы на внешние коммутационные структуры**

Количество внешних сетевых интерфейсов интегрированного ИСП ограничивается не только достижимой степенью интеграции СнК (системы на кристалле), но и допустимым количеством выводов корпуса микросхемы. Поэтому важным направлением развития ИСП будут высокоскоростные интерфейсы для подключения к внешним коммутационным структурам, позволяющим тесно соединять несколько ИСП и создавать многопортовые высокопроизводительные коммутаторы и маршрутизаторы. В этом плане перспективными представляются локальные последовательные интерфейсы Ethernet и SRIO.

Очевидное достоинство локального Ethernet со скоростями передачи 40 и 100 Гбит/с в качестве интерфейса на внешние коммутационные структуры заключается в его полной логической совместимости с сетевыми интерфейсами Ethernet, что позволяет унифицированно обслуживать все внешние пакетные интерфейсы ИСП. К сожалению, на сегодняшний день таких готовых решений нет, имеется лишь стандарт интерфейса объединительных плат 40GBASE-KR4 и предполагается появление стандарта на скорость 100 Гбит/с.

При более скромных требованиях, на уровне 6...24 Гбит/с на порт, уже сегодня готов к использованию для подключения к ИСП внешних коммутационных структур интерфейс SRIO. В отличие от компьютерных интерфейсов типа USB или PCIe, предлагающих централизованное управление и древовидную топологию, организация общения всех абонентов интерфейса SRIO между собой как равных с равными (peer-to-peer) в большей мере соответствует сетевым принципам взаимодействия через внешнюю коммутационную структуру [12]. Спецификация SRIO (части 10 и 11) предусматривает потоковые (streaming) транзакции с инкапсуляцией произвольных протоколов и элементами качества обслуживания [13], а имеющиеся на рынке интегральные коммутаторы SRIO обеспечивают суммарную пропускную способность до 240 Гбит/с [14].

## **7 Заключение**

Различные ИСП становятся основой все большего числа разнообразных сетевых устройств. По мере развития в них постоянно увеличивается число про-

граммируемых процессорных ядер, повышаются скорости сетевых интерфейсов, расширяется номенклатура различного рода аппаратных ускорителей. Процессорные ядра проявляют тенденцию к кластеризации, а кластеры — к функциональной специализации. Концептуально все эти компоненты ИСП включаются в виртуальный конвейер, по ступеням которого последовательно проходят обрабатываемые пакеты.

Сбалансированный виртуальный конвейер, у которого все ступени имеют примерно равную пропускную способность, обеспечивает детерминистическую обработку с гарантированной производительностью. Но на практике непросто сбалансировать несколько виртуальных конвейеров из одних и тех же блоков в рамках одного ИСП, поэтому особенное значение приобретает интеллигентный диспетчер очередей, регулирующий потоки пакетов внутри ИСП и обеспечивающий максимально возможную эффективность использования наличных ресурсов.

Пример компании Cisco Systems с ее новейшим масштабируемым ИСП nPower [15] показывает, что многоядерные ИСП уже готовы к преодолению рубежа производительности терabit в секунду.

## Литература

1. Шаршаков А. Реализация и развитие технологии ATM корпорацией IBM // Сети / Network World, 1998. № 1. <http://www.osp.ru/nets/1998/01/143040>.
2. Вейс Р. Сетевые кристаллы — основа телефонных и информационных сетей нового поколения // Мир компьютерной автоматизации, январь 2000. <http://asutp.ru/?p=600085>.
3. T4240: QorIQ T Series T4240/T4160 24/16 Virtual Core Communications Processors. [http://www.freescale.com/webapp/sps/site/prod\\_summary.jsp?code=T4240&nodeId=018rH325E40180](http://www.freescale.com/webapp/sps/site/prod_summary.jsp?code=T4240&nodeId=018rH325E40180).
4. OCTEON III CN7XXX Multi-Core MIPS64 Processors. [http://cavium.com/OCTEON-III\\_CN7XXX.html](http://cavium.com/OCTEON-III_CN7XXX.html).
5. Broadcom introduces industry's First 100 Gbps Full Duplex Network Processor. <http://www.broadcom.com/press/release.php?id=s666869>.
6. Беляев А., Путря Ф., Солохина Т., Юдинцев В. Многоядерные процессоры для устройств связи. Перспективы и проблемы // Электроника НТБ, 2011. Вып. 8. С. 90–104.
7. Layerscape architecture. A look inside the next-generation QorIQ LS Series SoCs. [http://cache.freescale.com/files/32bit/doc/white\\_paper/LAYARCHTECHOVWP.pdf](http://cache.freescale.com/files/32bit/doc/white_paper/LAYARCHTECHOVWP.pdf).
8. Егоров В. Б. Интегрированные сетевые процессоры Freescale с ARM-ядрами // Электронные компоненты, 2013. № 4. С. 85–89.
9. Егоров В. Б. Архитектурные инновации в многоядерных ИКМ QorIQ // Электронные компоненты, 2010. № 10. С. 66–72.
10. Егоров В. Б. Второе поколение интегрированных коммуникационных микроконтроллеров QorIQ от Freescale // Электронные компоненты, 2012. № 9. С. 99–103.

11. Егоров В. Б. Коммуникационные процессоры Axxia от LSI // Электронные компоненты, 2014. № 4. С. 78–83.
12. Егоров В. Б. Последовательный интерфейс RapidIO и его применение в пакетной коммутации // Электронные компоненты, 2008. № 12. С. 69–76.
13. RapidIO Specifications. <http://www.rapidio.org/specs/current>.
14. RapidIO Switches. <http://www.idt.com/products/interface-connectivity/serial-rapidio-solutions/rapidio-switches>.
15. Cisco Unveils nPower, World's Most Advanced Network Processor. <http://newsroom.cisco.com/release/1262342>.

Поступила в редакцию 17.07.14

---

## MODERN TRENDS IN EVOLUTION OF INTEGRATED NETWORK PROCESSOR ARCHITECTURES

V. Egorov

Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The article points out the increasing role of programmable integrated network processors (INP) in developments of up-to-day packet switches, routers, and various devices for network infrastructure. The article gives an analysis of recent trends in INP architecture evolution and distinguishes different programmability levels and the most significant INP macroarchitecture components. The article reveals interdependence between macroarchitecture and opportunity to arrange inside an INP internal virtual pipelines for deterministic packet processing with high throughput and shows the crucial role of a queue manager in achieving efficient functioning of the virtual pipelines inside an INP. The article mentions the advantages of high-speed serial interfaces with interaction on the peer-to-peer basis as an attaching facility to a shared external switching fabric.

**Keywords:** integrated network processor; multicore processor; macroarchitecture; virtual pipeline; queue manager; high speed serial interfaces

**DOI:** 10.14357/08696527140306

## References

1. Sharshakov, A. 1998. Realizatsiya i razvitiye tekhnologii ATM korporatsiy IBM [Implementation and development of the ATM technology by IBM]. *Seti / Network World* 1. Available at: <http://www.osp.ru/nets/1998/01/143040> (accessed July 17, 2014).
2. Veys, R. 2000. Setevye kristally — osnova telefonnykh i informatsionnykh setey novogo pokoleniya [Network chips are a base of telephone and information new generation networks]. January 2000. *Mir kompyuternoy avtomatizatsii* [Computer automation world]. Available at: <http://asutp.ru/?p=600085> (accessed July 17, 2014).

3. T4240: QorIQ T Series T4240/T4160 24/16 Virtual Core Communications Processors. Available at: [http://www.freescale.com/webapp/sps/site/prod\\_summary.jsp?code=T4240&nodeId=018rH325E40180](http://www.freescale.com/webapp/sps/site/prod_summary.jsp?code=T4240&nodeId=018rH325E40180) (accessed July 17, 2014).
4. OCTEON III CN7XXX Multi-Core MIPS64 Processors. Available at: [http://cavium.com/OCTEON-III\\_CN7XXX.html](http://cavium.com/OCTEON-III_CN7XXX.html) (accessed July 17, 2014).
5. Broadcom introduces industry's First 100 Gbps Full Duplex Network Processor. Available at: <http://www.broadcom.com/press/release.php?id=s666869> (accessed July 17, 2014).
6. *Belyaev, A., F. Putrya, T. Solokhina, and V. Yudintsev.* 2011. Mnogoyadernye protsessory dlya ustroystv svyazi. Perspektivy i problemy [Multicore processors for communication devices. Trends and problems]. *Elektronika NTB* [NTB Electronics] 8:90–104.
7. Layerscape architecture. A look inside the next-generation QorIQ LS Series SoCs. Available at: [http://cache.freescale.com/files/32bit/doc/white\\_paper/LAYARCHTECHOVWP.pdf/](http://cache.freescale.com/files/32bit/doc/white_paper/LAYARCHTECHOVWP.pdf/) (accessed July 17, 2014).
8. Egorov, V. 2013. Integrirovannyе protsessory Freescale s ARM-yadrami [Integrated network processors from Freescale with ARM cores]. *Elektronnye Komponenty* [Electronic Components] 4:85–89.
9. Egorov, V. 2010. Arkhitekturnye innovatsii v mnogoyadernykh IKM QorIQ ot Freescale [Architecture innovations in multicore ICM QorIQ from Freescale]. *Elektronnye Komponenty* [Electronic Components] 10:66–72.
10. Egorov, V. 2012. Vtoroe pokolenie integrirovannykh kommunikatsionnykh mikrokontollerov QorIQ ot Freescale [The second generation of the integrated communication controllers QorIQ from Freescale]. *Elektronnye Komponenty* [Electronic Components] 9:99–103.
11. Egorov, V. 2014. Kommunikatsionnye protsessory Axxia from LSI [Communication processors Axxia from LSI]. *Elektronnye Komponenty* [Electronic Components] 4:78–83.
12. Egorov, V. 2008. Posledovatelnyy interfeys RapidIO i ego primenenie v paketnoy kommutatsii [The serial interface RapidIO and its application in packet switching]. *Elektronnye Komponenty* [Electronic Components] 2:69–76.
13. RapidIO Specifications. Available at: <http://www.rapidio.org/specs/current> (accessed July 17, 2014).
14. RapidIO Switches. Available at: <http://www.idt.com/products/interface-connectivity/serial-rapidio-solutions/rapidio-switches> (accessed July 17, 2014).
15. Cisco Unveils nPower, World's Most Advanced Network Processor. Available at: <http://newsroom.cisco.com/release/1262342> (accessed July 17, 2014).

Received July 17, 2014

## Contributor

**Egorov Vladimir B.** (b. 1948) — Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problem, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; [vegorov@ipiran.ru](mailto:vegorov@ipiran.ru)

**ДВУХПАРАМЕТРИЧЕСКИЙ АНАЛИЗ  
МАГНИТНО-РЕЗОНАНСНОГО ИЗОБРАЖЕНИЯ МЕТОДОМ  
МАКСИМУМА ПРАВДОПОДОБИЯ В СРАВНЕНИИ  
С ОДНОПАРАМЕТРИЧЕСКИМ ПРИБЛИЖЕНИЕМ**

*Т. В. Яковлева<sup>1</sup>, Н. С. Кульберг<sup>2</sup>*

**Аннотация:** Исследуется метод двухпараметрического анализа данных магнитно-резонансного изображения, который позволяет получить одновременную оценку полезного сигнала и шума в анализируемом изображении на основе принципа максимума правдоподобия. Данный метод представляет собой принципиально новый способ обработки данных в условиях распределения Райса и может эффективно использоваться в составе информационных технологий для решения широкого круга задач, связанных с фильтрацией рисковых сигналов. Решение двухпараметрической задачи основывается исключительно на данных выборочных измерений и не связано с априорными предположениями относительно величины шума, которые неизбежно ограничивают точность однопараметрического приближения. Посредством компьютерного моделирования проведен сравнительный анализ традиционного однопараметрического и развитого двухпараметрического методов решения задачи оценивания полезной составляющей в величине сигнала, формирующего магнитно-резонансное изображение, рассчитаны статистические характеристики смещения и разброса искомых параметров сигнала и шума. Показаны преимущества двухпараметрического метода анализа данных.

**Ключевые слова:** распределение Райса; функция правдоподобия; метод максимума правдоподобия; дисперсия шума; отношение сигнала к шуму

**DOI:** 10.14357/08696527140307

## 1 Введение

Статистические методы при решении задач анализа и обработки случайных сигналов давно и успешно используются в самых различных областях научных исследований и практических приложений. В настоящей работе исследуется задача анализа и обработки магнитно-резонансных изображений как характерного

---

<sup>1</sup>Вычислительный центр имени А. А. Дородницына Российской академии наук,  
yakovleva@ccas.ru

<sup>2</sup>Вычислительный центр имени А. А. Дородницына Российской академии наук,  
kulberg@yandex.ru

примера применимости статистической модели Райса. Как известно, измеряемой и анализируемой величиной при этом является амплитуда или огибающая сигнала, формирующего магнитно-резонансное изображение и представляющего собой сумму искомого исходного сигнала и случайного шума, образованного многими независимыми нормально-распределенными слагаемыми с нулевым средним значением. Величина амплитуды такого суммарного сигнала подчиняется статистическому распределению Райса [1, 2]. Круг задач, математически описываемых статистической моделью Райса, помимо задач магнитно-резонансной визуализации [3–8] охватывает такие задачи, как прием и обработка радиосигналов [9]; анализ сигналов радара [10, 11]; задачи, связанные с анализом звукового эхосигнала [12, 13]; определение характеристик оптической среды путем изучения свойств отраженного периодически модулированного света [14].

Традиционные методы анализа и обработки райсовских сигналов в задачах магнитно-резонансной визуализации, основанные на принципе максимума правдоподобия, как правило, предполагают оценивание лишь одного из двух статистических параметров — параметра средней величины сигнала в предположении, что второй параметр — дисперсия шума — является известным априори (см., например, [6]). Эта так называемая однопараметрическая модель, используемая многими авторами при решении задачи, никогда не реализуется на практике, и поэтому ее применение является серьезным ограничением традиционных подходов к обработке райсовских сигналов. В работе [15] впервые была рассмотрена адекватная двухпараметрическая модель для решения задачи, т. е. модель, предполагающая наличие двух неизвестных параметров анализируемого райсовского сигнала: средней величины сигнала и дисперсии шума. В [15] предпринята попытка совместной оценки обоих указанных статистических параметров методом максимума правдоподобия. Однако решение двухпараметрической задачи сопряжено со значительными трудностями как теоретического, так и вычислительного характера, поскольку приходится рассматривать систему двух существенно нелинейных уравнений. Отчасти поэтому до недавнего времени теоретическое изучение задачи ограничивалось лишь оценками стандартного отклонения по методу Крамера–Рао [15] и предположениями относительно свойств решения задачи, основанными на графических иллюстрациях, а не на строгом математическом анализе [7].

Двухпараметрический метод решения задачи [16–18] обеспечивает одновременное определение обоих априори не известных статистических параметров — средней величины полезного сигнала и дисперсии шума. Эти два параметра случайной величины, формирующей магнитно-резонансное изображение, позволяют, в свою очередь, полностью определить функцию, описывающую восстанавливаемое исходное, т. е. незашумленное изображение. Искомые параметры вычисляются лишь на основании выборок измерений, без каких-либо ограничений, налагаемых предположением об априорной известности шума, и поэтому двухпараметрический метод обеспечивает гораздо более корректные оценки по сравнению с однопараметрическим приближением. Математическое обоснование

двуихпараметрического метода, основанное на строгом доказательстве существования и единственности решения задачи, впервые представлено в работах [16–18].

Развиваемый двухпараметрический метод можно обобщенно рассматривать как инструмент анализа состояний стохастической системы в случае, когда система характеризуется двумя состояниями: исходным незашумленным состоянием и состоянием системы, подверженной воздействию на нее гауссовского шума. Задача фильтрации магнитно-резонансного изображения состоит в том, чтобы на основе выборочных данных о «зашумленном» состоянии системы восстановить исходное «незашумленное» состояние, соответствующее искомым данным, анализируемым посредством изучения сформированного ими магнитно-резонансного изображения. С этой точки зрения двухпараметрический метод восстановления исходного, незашумленного магнитно-резонансного изображения представляет собой новый метод фильтрации данных, который может использоваться в составе соответствующей информационной технологии.

Задача, решаемая в настоящей работе, состоит в изучении особенностей и статистических характеристик двухпараметрического метода анализа магнитно-резонансного изображения, в частности путем сопоставления с однопараметрическим приближением. В работе рассматриваются преимущества и ограничения, характерные для обоих методов, выявляются условия их оптимизации. Основные результаты работы, полученные посредством численного моделирования сравниваемых методов, представлены в разд. 3, в то время как в разд. 2 кратко изложены теоретические основы двухпараметрического анализа данных в условиях распределения Райса. Теория двухпараметрического подхода к анализу и обработке райсовских данных развита Т. В. Яковлевой. В частности, в результате проведенного теоретического исследования разработаны алгоритмы расчета неизвестных параметров. Данные алгоритмы легли в основу созданных Н. С. Кульбергом программ расчета искомых параметров задачи. На основе разработанного программного обеспечения [19] в работе проведены численные эксперименты с целью сопоставления одно- и двухпараметрической моделей.

## **2 Теоретические основы двухпараметрического анализа данных методом максимума правдоподобия в условиях распределения Райса**

Рассмотрим задачу анализа случайного сигнала, представляющего собой модуль комплексной величины, действительная  $x_{\text{Re}}$  и мнимая  $x_{\text{Im}}$  части которого искажаются гауссовским шумом, имеющим нормальное распределение с некоторой неизвестной величиной дисперсии и нулевыми средними значениями. Такие сигналы характеризуют, в частности, процесс построения магнитно-резонансного изображения. Амплитуда  $x = \sqrt{x_{\text{Re}}^2 + x_{\text{Im}}^2}$  формирует анализируемое магнитно-резонансное изображение. Будем использовать следующие обозначения:  $\nu$  — математическое ожидание действительной и мнимой частей измеряемого сигнала;  $\sigma^2$  — величина дисперсии гауссовского шума, искажающего сигнал. Величи-

на  $\nu/\sigma$  характеризует отношение сигнала к шуму. Тогда амплитуда  $x$  имеет распределение Райса с плотностью распределения вероятности

$$P(x|\nu, \sigma^2) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2 + \nu^2}{2\sigma^2}\right) I_0\left(\frac{x\nu}{\sigma^2}\right). \quad (1)$$

Здесь и ниже будем использовать следующие обозначения:  $I_\alpha(z)$  — модифицированная функция Бесселя первого рода (или функция Инфельда) порядка  $\alpha$ ;  $x_i$  — величина сигнала  $i$ -й выборки при измерении данных для последующей обработки изображения;  $n$  — количество элементов в выборке (длина выборки). Для обозначения различных видов усреднения по выборке будем использовать угловые скобки, а среднее значение при бесконечно большой длине выборки будем обозначать чертой сверху:

$$\begin{aligned} \langle x \rangle &= \frac{1}{n} \sum_{i=1}^n x_i; & \langle x^2 \rangle &= \frac{1}{n} \sum_{i=1}^n x_i^2; & \dots; & \langle x^k \rangle &= \frac{1}{n} \sum_{i=1}^n x_i^k; \\ \overline{x} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n x_i; & \overline{x^2} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n x_i^2; & \dots; & \overline{x^k} &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n x_i^k. \end{aligned}$$

Функция правдоподобия, как известно, представляет собой плотность совместного распределения в выборке как функцию статистических параметров  $\nu$  и  $\sigma^2$  и определяется произведением функций плотности вероятности для каждого измерения из выборки:

$$L(\nu, \sigma^2) = \prod_{i=1}^n P(x_i|\nu, \sigma^2). \quad (2)$$

Для решения задачи методом максимума правдоподобия удобно использовать логарифмическую функцию правдоподобия  $\ln L(\nu, \sigma^2)$  [6, 16–18]. Опуская промежуточные выкладки [16–18], приведем основные соотношения, определяющие алгоритм вычисления искомых параметров  $\nu$  и  $\sigma^2$  двухпараметрическим методом. Система уравнений максимума правдоподобия для  $\nu$  и  $\sigma^2$  имеет следующий вид [16–18]:

$$\left. \begin{aligned} \nu - \frac{1}{n} \sum_{i=1}^n x_i \tilde{I}\left(\frac{x_i \nu}{\sigma^2}\right) &= 0; \\ \sigma^2 - \frac{1}{2n} \sum_{i=1}^n (x_i^2 + \nu^2) + \frac{\nu}{n} \sum_{i=1}^n x_i \tilde{I}\left(\frac{x_i \nu}{\sigma^2}\right) &= 0. \end{aligned} \right\} \quad (3)$$

В уравнениях (3) используется обозначение  $\tilde{I}(z) = I_1(z)/I_0(z)$ . Свойства функции  $\tilde{I}(z)$  подробно изучены в предыдущих работах авторов (см., например, [18]). Именно эти свойства: гладкость, монотонность, ограниченность,

выпуклость — определяют существование и единственность решения системы уравнений (3), строго доказанные в [18]. Решение системы уравнений (3) максимума правдоподобия на основе результатов выборки из  $n$  измерений величины сигнала  $x_1, \dots, x_n$  позволяет получить наиболее вероятные значения параметров  $\nu$  и  $\sigma^2$  случайного сигнала, формирующего изображение, и тем самым решить задачу восстановления неискаженного шумами исходного изображения.

Решение первого уравнения системы (3) независимо от второго означает решение однопараметрической задачи определения одного неизвестного параметра  $\nu$  в предположении, что второй параметр — дисперсия  $\sigma^2$  — является известным априори [6]. Существование и единственность решения однопараметрической задачи в предположении, что дисперсия шума  $\sigma^2$  известна априори, строго доказаны в [16, 18]. Развитие метода максимума правдоподобия на случай, когда оба статистических параметра  $\nu$  и  $\sigma^2$  априори не известны, приводит к необходимости решения системы нелинейных уравнений (3). В результате ряда преобразований, на которых здесь останавливаться не будем, система (3) для  $\nu$  и  $\sigma^2$  может быть представлена в виде:

$$\left. \begin{aligned} \nu &= \frac{1}{n} \sum_{i=1}^n x_i \tilde{I} \left( \frac{2x_i \nu}{\langle x^2 \rangle - \nu^2} \right); \\ \sigma^2 &= \frac{1}{2} (\langle x^2 \rangle - \nu^2). \end{aligned} \right\} \quad (4)$$

Первое из уравнений системы (4) — уравнение для  $\nu$  — представляет собой уравнение с одним неизвестным, т. е. задача решения системы уравнений (3) для двух неизвестных  $\nu$  и  $\sigma^2$  сведена к решению одного уравнения для одной переменной  $\nu$  (первое уравнение системы (4)) и последующего расчета второго неизвестного по заданной формуле (второе уравнение системы (4)). Следовательно, решение двухпараметрической задачи существенно упрощается, а объем необходимых вычислительных ресурсов практически не увеличивается по сравнению с однопараметрическим методом. Как показано в [16–18], решение системы уравнений (3) (или (4)) двухпараметрического метода максимума правдоподобия существует и является единственным. Представленные в следующем разделе результаты численных экспериментов получены решением первого из уравнений (3) и решением системы (4) для однопараметрического и двухпараметрического метода соответственно.

### 3 Результаты численного моделирования

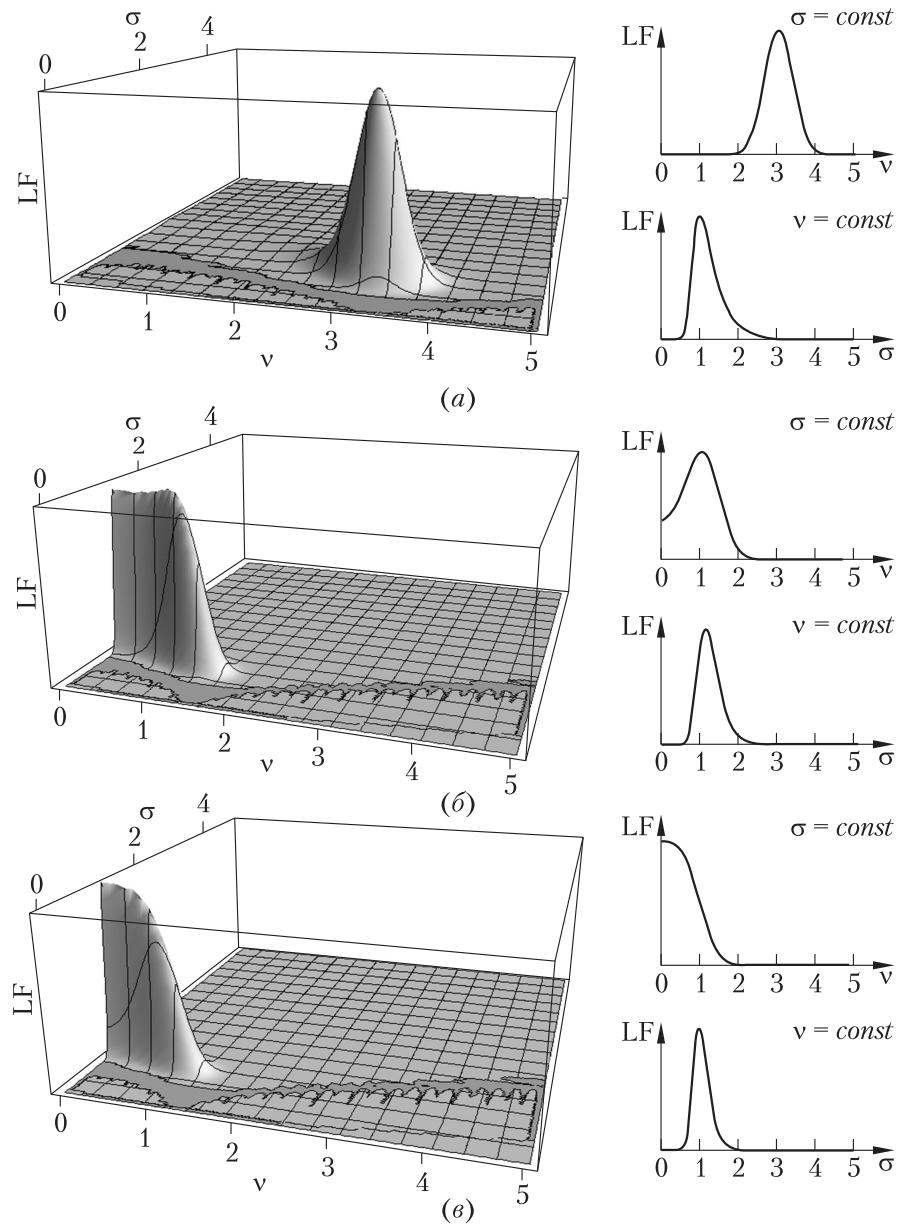
В данном разделе приведены результаты численного моделирования изложенных выше одно- и двухпараметрического методов фильтрации магнитно-резонансного изображения посредством решения задачи расчета искомых параметров

райсовского распределения на основе принципа максимума правдоподобия. Очевидно, что свойства решений уравнений, соответствующих обоим сопоставляемым методам, определяются особенностями функции правдоподобия распределения Райса [17]. Чтобы проиллюстрировать эти свойства, приведем трехмерные графики данной функции, полученные в результате численных экспериментов путем ее расчета по формуле (2) по выборочным данным, генерируемым датчиком случайных чисел с распределением Райса.

На рис. 1 приведены полученные посредством численного моделирования трехмерные графики функции правдоподобия как функции двух переменных: математического ожидания и дисперсии шума, — построенные для различных значений отношения сигнала к шуму, которые наглядно иллюстрируют существование и единственность максимума функции правдоподобия, или, другими словами, единственность решения двухпараметрической задачи методом максимума правдоподобия, что, как отмечалось выше, декларируется теоремами, доказанными в [16, 18]. На рис. 1 представлены также сечения функции правдоподобия плоскостью  $\sigma = \text{const}$  и плоскостью  $\nu = \text{const}$  в окрестности точки максимума. Расчет функции правдоподобия производился по формуле (2). При этом использовался датчик случайных чисел, подчиняющихся распределению Райса с соответствующими параметрами: (а)  $\nu = 3, \sigma = 1$ ; (б)  $\nu = 1, \sigma = 1$ ; (в)  $\nu = 0,3, \sigma = 1$ . Длина выборки  $n$  во всех вариантах расчета функции правдоподобия, представленных на рис. 1, составляла 8.

Помимо наличия единственной точки максимума функции правдоподобия рис. 1 иллюстрирует также тот факт, что с ростом отношения сигнала к шуму  $\nu/\sigma$  максимум функции правдоподобия приобретает более острую форму и, соответственно, график сечения данной функции как плоскостью  $\sigma = \text{const}$ , так и плоскостью  $\nu = \text{const}$  сужается.

Другая важная особенность формы функции правдоподобия состоит в следующем. Очевидно, что в результате сечения трехмерной поверхности данной функции вблизи точки ее максимума плоскостями  $\sigma = \text{const}$  (верхние графики справа от трехмерного изображения) и плоскостями  $\nu = \text{const}$  (нижние графики справа от трехмерного изображения) получаются функции, которые характеризуют зависимость функции правдоподобия от каждого из параметров  $\nu$  и  $\sigma$  соответственно при фиксированной величине второго параметра. При этом из представленных на рис. 1 графиков следует, что характерная ширина функции, описывающей зависимость функции правдоподобия от параметра сигнала  $\nu$  (при  $\sigma = \text{const}$ ), заметно больше, чем характерная ширина функции, которая отражает зависимость функции правдоподобия от параметра стандартного отклонения  $\sigma$  (при  $\nu = \text{const}$ ). Это вполне закономерно в силу того, что параметры  $\nu$  и  $\sigma$  входят в функцию правдоподобия (2) как ее аргументы по-разному, поскольку они по-разному входят в соответствующую распределению Райса функцию плотности вероятности (1). Это различие в зависимостях функции правдоподобия от каждого из своих аргументов  $\nu$  и  $\sigma$  влечет за собой важное следствие: точность определения параметров  $\nu$  и  $\sigma$  методом максимума правдоподобия будет раз-



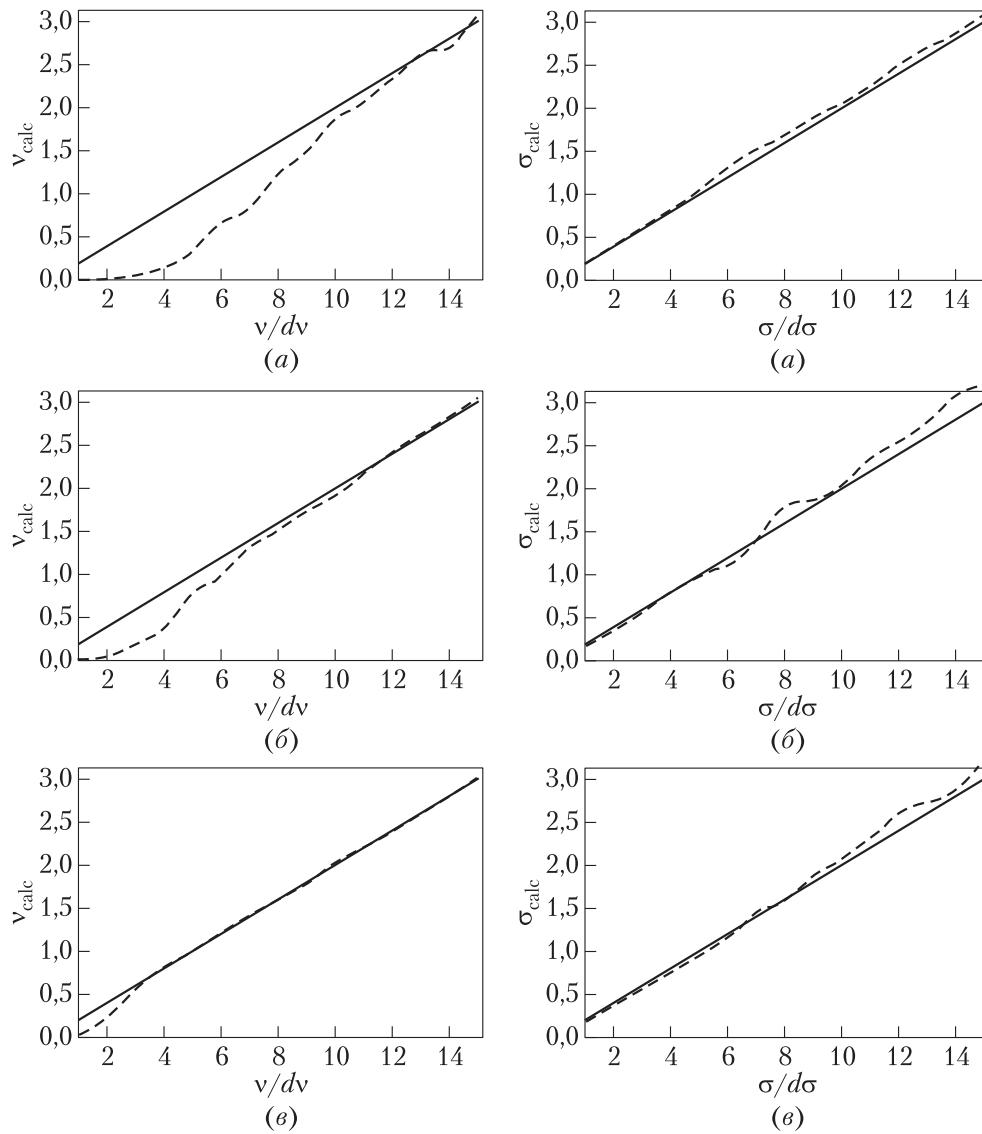
**Рис. 1** Функция правдоподобия и ее сечения при различных соотношениях параметров  $\nu$ ,  $\sigma$ : (а)  $\nu = 3$ ,  $\sigma = 1$ ,  $\nu/\sigma = 3$ ; (б)  $\nu = 1$ ,  $\sigma = 1$ ,  $\nu/\sigma = 1$ ; (в)  $\nu = 3$ ,  $\sigma = 1$ ,  $\nu/\sigma = 0,3$

личной, а именно: точность определения  $\sigma$  как аргумента более узкой функции правдоподобия в сечении  $\nu = const$  должна быть выше, чем точность определения  $\nu$  как аргумента с большей характерной шириной функции правдоподобия в сечении  $\sigma = const$ . Этот важный вывод подтвержден результатами численных экспериментов и проиллюстрирован на рис. 2 и 3.

На рис. 2 и 3 представлены результаты численного расчета искомых параметров  $\nu$  и  $\sigma$  посредством двухпараметрического метода максимума правдоподобия для различных значений отношения сигнала к шуму с целью более точного выявления границ применимости данного метода. Численный эксперимент проводился следующим образом. Генерировались данные, подчиняющиеся распределению Райса, для различных исходно заданных значений параметров сигнала: средней величины  $\nu$  и стандартного отклонения  $\sigma$ . Используя данные выборок значений сигнала  $x_i$  ( $i = 1, \dots, n$ ), рассчитывались вышеуказанные статистические параметры на основе разработанного двухпараметрического алгоритма: путем решения первого из уравнений максимума правдоподобия (4) для определения  $\nu$  и затем вычисления параметра  $\sigma$  на основе второго уравнения (4). Полученные расчетные значения параметров  $\nu$  и  $\sigma$  представлены на рис. 2 и 3: расчетные значения для параметра  $\nu$  представлены на рис. 2, *a–2, в*, а значения для параметра  $\sigma$  представлены на рис. 3, *a–3, в*. Графики на рис. 2 и 3 соответствуют следующим условиям: исходные значения обоих параметров  $\nu$  и  $\sigma$  изменялись в диапазоне от 0,2 до 3,0 с шагом 0,2 (по оси абсцисс отмечены точки отсчета). Длина выборки  $n$  составляла 16, вычисления усреднялись по 50 выборкам.

На рис. 2, *a–2, в* показаны зависимости отклонения расчетных значений  $\nu$  (штриховые кривые) от реальной, исходно заданной величины этого параметра (сплошная прямая линия) при различных исходных значениях параметра  $\sigma$ , т. е. в различных сечениях  $\sigma = const$ . Отклонение штриховой кривой от сплошной прямой характеризуют точность расчетов. При переходе к каждому последующему графику от рис. 2, *a* до 2, *в* отношение сигнала к шуму  $\nu/\sigma$  увеличивается обратно пропорционально  $\sigma$ , при этом точность расчетов тоже увеличивается. Кроме того, зависимость точности от отношения  $\nu/\sigma$  можно проследить на каждом из представленных графиков на рис. 2, *a–2, в*: при перемещении слева направо вдоль расчетных кривых, т. е. с ростом параметра  $\nu$ , происходит увеличение отношения сигнала к шуму и точность расчетов заметно растет.

На рис. 3, *a–3, в* показаны зависимости отклонения расчетных значений параметра  $\sigma$  (штриховые кривые) от реальных, исходно заданных значений этого параметра (сплошные прямые линии) при различных исходных значениях параметра  $\nu$ , т. е. в различных сечениях  $\nu = const$ . Отклонение штриховой кривой от сплошной прямой на обоих графиках характеризуют точность расчетов. При переходе к каждому последующему графику на рис. 3, *a–3, в* отношение сигнала к шуму растет пропорционально величине параметра  $\nu$ . Как и в случае зависимостей, представленных на рис. 2, графики на рис. 3 иллюстрируют зависимость точности расчетов от значения отношения сигнала к шуму: с ростом параметра  $\sigma$



**Рис. 2** Результаты численного расчета параметра  $\nu$  посредством двухпараметрического метода максимума правдоподобия. На графиках  $\nu$  изменяется в диапазоне от 0,2 до 3,0 с шагом 0,2,  $n = 16$ : (a)  $\sigma = 1,2$ ; (б) 0,8; (в)  $\sigma = 0,4$

**Рис. 3** Результаты численного расчета параметра  $\sigma$  посредством двухпараметрического метода максимума правдоподобия. На графиках  $\sigma$  изменяется в диапазоне от 0,2 до 3,0 с шагом 0,2,  $n = 16$ : (а)  $\nu = 1$ ; (б) 2; (в)  $\sigma = 3$

(при перемещении слева направо вдоль расчетных кривых) происходит снижение отношения сигнала к шуму и точность расчетов заметно падает.

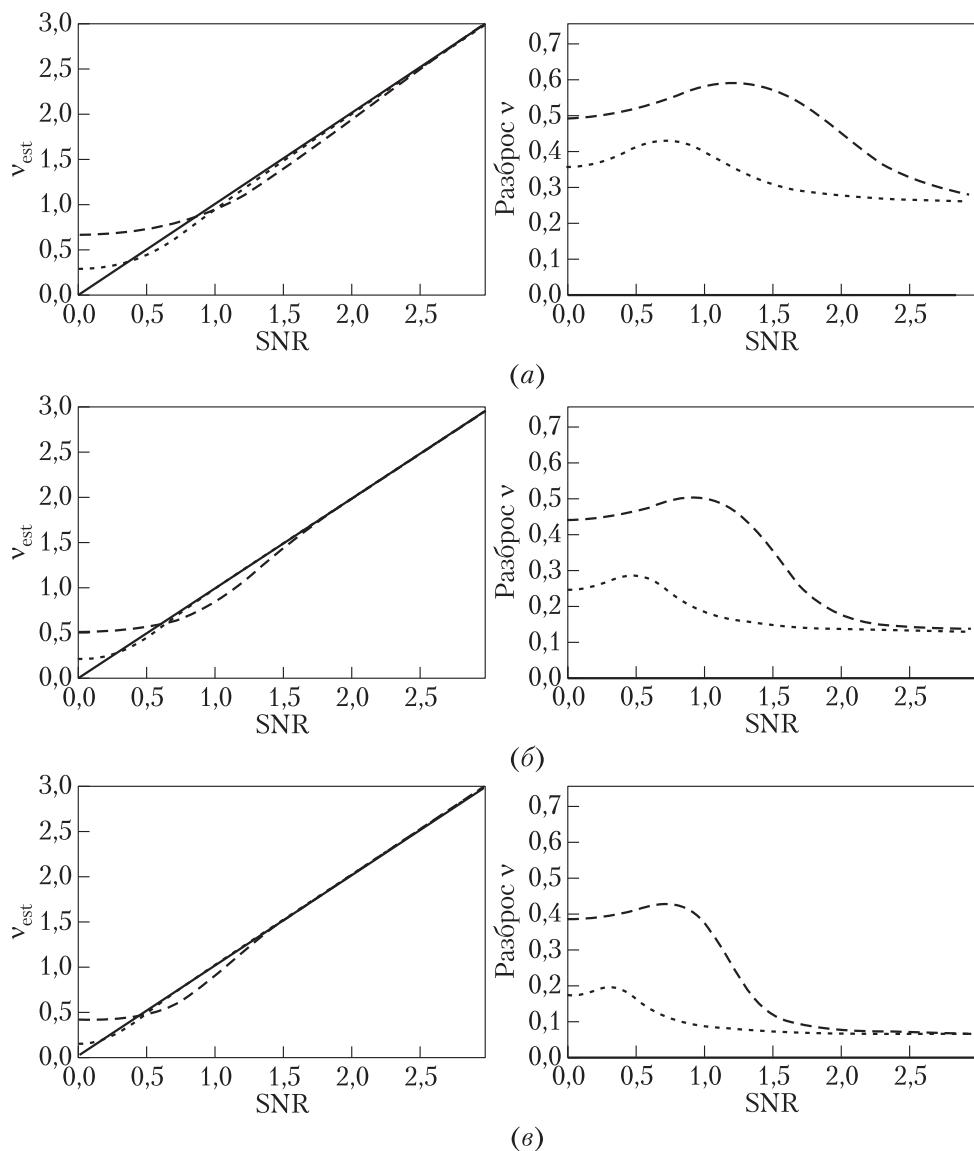
Повышение точности расчетов параметров  $\nu$  и  $\sigma$  методом максимума правдоподобия с ростом отношения сигнала к шуму обусловлено тем, что с ростом  $\nu/\sigma$  максимум функции правдоподобия становится более выраженным (см. рис. 1), приобретает более острую форму, что влечет за собой повышение точности расчета соответствующих данному максимуму значений  $\nu$  и  $\sigma$ .

Другой важный вывод, следующий из графиков, представленных на рис. 2 и 3, состоит в том, что точность оценки параметра средней величины сигнала  $\nu$  посредством двухпараметрического метода максимума правдоподобия заметно ниже, чем точность оценки параметра стандартного отклонения  $\sigma$ . Кроме того, как следует из графиков на рис. 2, точность метода существенно повышается начиная с некоторого граничного значения отношения сигнала к шуму  $\nu/\sigma$ , причем для оценки параметра  $\nu$  соответствующий данному граничному значению диапазон относительно высокой точности определяется соотношением  $\nu/\sigma \geq 2$ , в то время как для оценки параметра  $\sigma$  диапазон высокой точности расчетов несколько больше:  $\nu/\sigma \geq 1,5$ . Этот вывод не является неожиданным: он следует из особенностей формы функции правдоподобия как функции параметров  $\nu$  и  $\sigma$ , которые отмечались в комментариях к рис. 1: характерная ширина функции правдоподобия как функции параметра  $\nu$  (в сечениях  $\sigma = const$ ) заметно больше, чем характерная ширина функции правдоподобия как функции параметра стандартного отклонения  $\sigma$  (при  $\nu = const$ ). Это различие влечет за собой различие в точности определения этих параметров методом максимума правдоподобия, а именно: точность определения  $\sigma$  как аргумента более узкой функции в сечении  $\nu = const$  превышает точность определения  $\nu$  как аргумента с большей характерной шириной функции в сечении  $\sigma = const$ , что и подтверждается графиками на рис. 2 и 3.

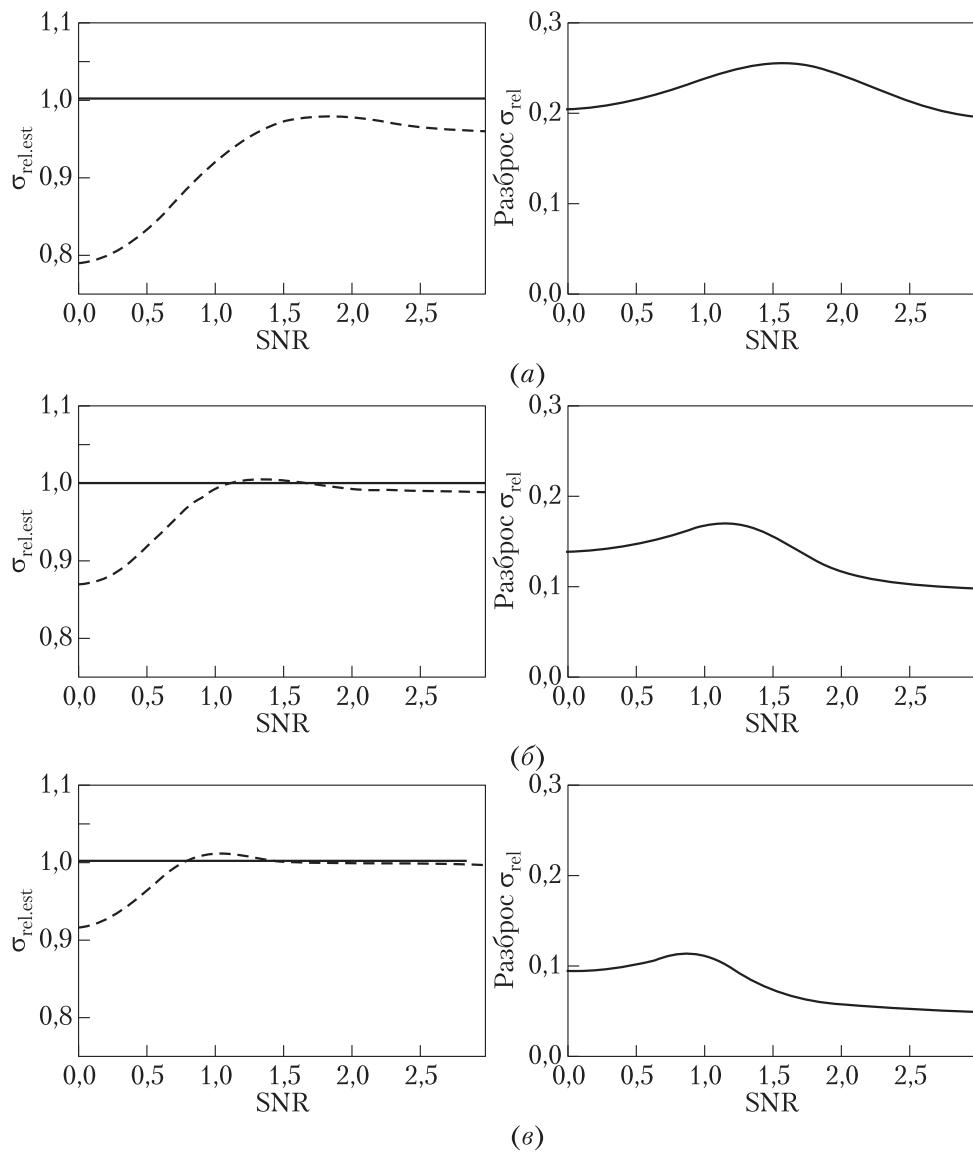
С целью сопоставления точности однопараметрического и двухпараметрического методов было проведено компьютерное моделирование задачи анализа рэйсовского сигнала, формирующего магнитно-резонансное изображение, указанными методами.

На рис. 4 и 5 представлены результаты численных экспериментов, наиболее приближенных к реальной ситуации анализа и обработки данных магнитно-резонансного изображения. При этом происходит усреднение по большому числу выборок измерений.

Численные эксперименты моделировались для следующих условий. Данные выборок измерений формировались с использованием датчика случайных чисел, распределенных по Райсу с различными исходными значениями параметров. Таким образом формировались различные наборы данных, соответствующие выборкам измерений, по которым путем решения соответствующих уравнений вычислялись искомые параметры: в случае двухпараметрической задачи оба параметра рассчитывались путем решения системы уравнений (4), а расчет па-



**Рис. 4** Графики, характеризующие смещение (левый столбец) и разброс (правый столбец) данных при расчете параметра  $\nu$  однопараметрическим (пунктирные кривые) и двухпараметрическим (штриховые кривые) методом максимума правдоподобия: (а)  $n = 16$ ; (б)  $n = 64$ ; (в)  $n = 256$



**Рис. 5** Графики, характеризующие смещение (левый столбец) и разброс (правый столбец) данных при расчете параметра  $\sigma$  двухпараметрическим методом максимума правдоподобия: (а)  $n = 16$ ; (б) 64; (в)  $n = 256$

раметра  $\nu$  в однопараметрическом приближении осуществлялся путем решения первого уравнения системы (3).

Полученные в результате численного решения значения параметров сравнивались с исходно заданными значениями, в результате чего анализировалась точность решения задачи как в однопараметрическом, так и в двухпараметрическом случае.

На рис. 4 приведены графики, иллюстрирующие точность расчетов параметра  $\nu$  однопараметрическим (пунктирные кривые) и двухпараметрическим (штриховые кривые) методом. Численные эксперименты проводились при различных значениях длины выборки:  $n = 16, 64$  и  $256$  на рис. 4, *a*, 4, *b* и 4, *v* соответственно. Для всех  $\nu$  и  $n$  осуществлялось по  $10^5$  измерений. По оси абсцисс на рис. 4 отложена исходно заданная величина параметра  $\nu$  в диапазоне от 0 до 3, в то время как величина  $\sigma$  была принята равной 1, так что точки оси абсцисс фактически соответствуют величине отношения сигнала к шуму  $\text{SNR} = \nu/\sigma$ . В левом столбце рис. 4 приведены графики рассчитанных значений  $\nu$ . Отклонение кривых от прямой линии характеризует точность сопоставляемых однотипных параметрических методов. В правом столбце приведены графики среднеквадратичных отклонений полученных значений параметра  $\nu$  в зависимости от величины отношения сигнала к шуму.

Аналогичные данные, полученные в результате численного оценивания параметра  $\sigma$ , приведены на рис. 5. Естественно, что в данном случае речь идет только о двухпараметрическом методе. Численные эксперименты проводились при различных значениях длины выборки:  $n = 16, 64$  и  $256$  на рис. 5, *a*, 5, *b* и 5, *v* соответственно. Для всех  $\nu$  и  $n$  осуществлялось по  $10^5$  измерений. Графики, представленные на рис. 5, получены при тех же условиях эксперимента, которые указаны в комментариях к рис. 4. Горизонтальная линия соответствует исходно заданному значению  $\sigma = 1$ . В левом столбце рис. 5 приведены графики рассчитанных значений  $\sigma$ , а в правом столбце — графики среднеквадратичных отклонений для  $\sigma$  в зависимости от отношения сигнала к шуму.

Таким образом, графические результаты численных экспериментов, представленные на рис. 4 и 5, иллюстрируют следующие выводы:

- (1) точность расчетов искомых параметров методом максимума правдоподобия ожидаемо возрастает с увеличением отношения сигнала к шуму, а также с увеличением длины выборки  $n$ ;
- (2) разброс данных при расчете искомых параметров заметно уменьшается с увеличением отношения сигнала к шуму, а также с увеличением длины выборки  $n$ ;
- (3) точность расчетов параметра  $\nu$  однопараметрическим методом при малых значениях отношения сигнал/шум ( $\nu < 2\sigma$ ) превосходит точность расчета данного параметра двухпараметрическим методом, что вполне ожидаемо и объясняется следующим образом: в двухпараметрической задаче присутствуют две неизвестные величины — значения обоих искомых параметров  $\nu$

и  $\sigma$ , при этом на ошибку при оценке параметра  $\nu$  в двухпараметрической задаче неизбежно влияет и ошибка в оценке второго параметра  $\sigma$ , в то время как в однопараметрической задаче параметр  $\sigma$  рассматривается как известный заранее, т. е. на результат работы традиционного метода не оказывает влияния погрешность определения параметра  $\sigma$ . В этой связи целесообразно еще раз подчеркнуть, что случай единственного неизвестного параметра на практике никогда не реализуется, поскольку  $\sigma$  никогда нельзя считать известным априори, и поэтому иллюстрируемое рис. 4 сравнение является весьма условным;

- (4) при достаточно больших значениях сигнала ( $\nu \geq 2\sigma$ ) результат двухпараметрического метода по определению параметра сигнала  $\nu$  становится практически неотличимым от результата традиционного метода при тех же значениях  $n$ . Что касается определения второго параметра задачи — параметра стандартного отклонения  $\sigma$ , то точность его расчета существенно повышается уже при  $\nu \geq 1,5\sigma$ .

Важная особенность полученных численных результатов, состоящая в том, что как смещение, так и разброс значений вычисленного параметра математического ожидания  $\nu$  значительно растут с уменьшением величины отношения сигнала к шуму, характерна для обоих исследуемых вариантов оценки параметров методом максимума правдоподобия (одно- и двухпараметрического). Как отмечалось выше, это объясняется особенностями формы функции правдоподобия, показанной на рис. 1, а также графиков сечения этой функции плоскостями  $\sigma = const$ : чем меньше отношение сигнала к шуму, тем менее выражен максимум функции правдоподобия, хотя этот максимум всегда существует.

#### 4 Заключение

В работе рассмотрена задача двухпараметрического анализа магнитно-резонансного изображения методом максимума правдоподобия в сопоставлении с традиционно используемым однопараметрическим приближением. Показана адекватность двухпараметрической модели изучаемой стохастической системы, принципиальным отличием которой от традиционной однопараметрической модели является отсутствие ограничений, связанных с предположением об априорной известности величины шума. На основе двухпараметрической модели разработан метод одновременного расчета параметров сигнала и шума в условиях распределения Райса посредством использования принципа максимума правдоподобия. Несомненным преимуществом разработанного метода является то, что расчет искомых параметров основывается исключительно на данных выборочных измерений, в то время как точность однопараметрического подхода неизбежно ограничивается априорными предположениями относительно величины шума.

В работе проведено компьютерное моделирование одно- и двухпараметрического метода обработки райсовского сигнала, формирующего магнитно-ре-

зонансное изображение. Сопоставление результатов численных экспериментов показало, что при достаточно большой длине выборки двухпараметрический метод по точности не уступает однопараметрическому приближению.

Применительно к анализу изображений существенным моментом развитого двухпараметрического метода является отсутствие ограничений, присущих решениям задач шумоподавления и обработки изображений линейными методами, в частности методами, связанными с решениями уравнений в частных производных и методами вейвлет-преобразований. Поэтому рассматриваемый в работе метод обработки данных и изображений, основанный на принципе максимума правдоподобия, обеспечивает эффективное подавление в том числе и нелинейных искажений, к каковым и относятся райсовские шумы, характерные для магнитно-резонансных изображений.

Важным для практических применений результатом проведенного исследования является доказанная возможность определения двух искомых параметров: сигнала и шума — без привлечения каких-либо дополнительных вычислительных ресурсов, что определяет существенное преимущество двухпараметрического метода по сравнению с традиционно используемым однопараметрическим приближением. Это стало возможным в силу того, что теоретическое рассмотрение задачи позволило свести решение системы двух уравнений с двумя неизвестными к решению одного уравнения с одним неизвестным.

Значимым преимуществом развитого двухпараметрического метода как потенциальной составной части новой информационной технологии, используемой для фильтрации данных при магнитно-резонансной визуализации, является единственность решения задачи методом максимума правдоподобия. Это имеет существенное значение для практического применения: так, получение первого численного решения уравнения для параметра  $\nu$  означает возможность прерывания процедуры численного расчета, так как полученное решение является единственным. В отсутствие доказательства единственности это было бы невозможно и для решения задачи потребовались бы значительно большие ресурсы компьютерного времени и большие объемы вычислений.

Таким образом, исследованный в работе двухпараметрический метод максимума правдоподобия, лежащий в основе новой информационной технологии, представляет собой эффективный инструмент для анализа и обработки данных в условиях распределения Райса, в частности при решении задачи фильтрации магнитно-резонансного изображения, и представляет собой значимый шаг вперед по сравнению с традиционно используемым однопараметрическим приближением.

## **Литература**

1. Nakagami M. Study of resultant amplitude of many vibrations whose phases and amplitudes are at random // Nippon Elec. Commun. Eng., 1940. No. 22. P. 69–92.
2. Rice S. O. Mathematical analysis of random noise // Bell Syst. Tech. J., 1944. Vol. 23. P. 282–322.

3. Henkelman R. M. Measurement of signal intensities in the presence of noise in MR images // Med. Phys., 1985. Vol. 12. No. 2. P. 232–233.
4. Wang T., Lei T. Statistical analysis of MR imaging and its application in image modeling // IEEE Conference (International) Image Processing and Neural Networks Proceedings, 1994. Vol. I. P. 866–870.
5. Gudbjartsson H., Patz S. The Rician distribution of noisy MRI data // Magnet. Reson. Med., 1995. Vol. 34. P. 910–914.
6. Sijbers J., den Dekker A. J., Scheunders P., van Dyck D. Maximum-likelihood estimation of Rician distribution parameters // IEEE Trans. Med. Imaging, 1998. Vol. 17. No. 3. P. 357–361.
7. Carrobbi C. F. M., Cati M. The absolute maximum of the likelihood function of the Rice distribution: Existence and uniqueness // IEEE Trans. Instrum. Meas., 2008. Vol. 57. No. 4. P. 682–689.
8. Sheil W. C. Magnetic resonance imaging (MRI scan). MedicineNet.com. Retrieved April 27, 2012.
9. Первов А. И. Статистическая теория радиотехнических систем. — М.: Радиотехника, 2003. 400 с.
10. Marcum J. I., Swelling P. Studies of target detection by pulsed radar // IRE Trans. Inform., 1960. Vol. IT-6. P. 59–267.
11. Gardner L. A. Statistical methods for the analysis of reentry vehicle radar data. Part 1. Lexington: M.I.T. Lincoln Lab., January 13, 1966. Mass. Tech. Note 1966-5.
12. Talukdar K. K., Lawing W. D. Estimation of the parameters of Rice distribution // J. Acoust. Soc. Am., 1991. Vol. 89. No. 3. P. 1193–1197.
13. Stuber G. L. Principles of mobile communication. — 2nd ed. — Norwell, MA: Kluwer, 2001. 752 p.
14. Яковлева Т. В., Князьков А. В. Двухпараметрический метод моментов обработки результатов периодической модуляции отраженного света для оценки ЭО коэффициентов // Лазеры, измерения, информация-2014: Тр. конф. — СПб.: СПбГПУ, 2014. С. 39.
15. Benedict T. R., Soong T. T. The joint estimation of signal and noise from the sum envelope // IEEE Trans. Inform. Theory, 1967. Vol. IT-13. No. 3. P. 447–454.
16. Yakovleva T. V., Kulberg N. S. Noise and signal estimation in MRI: Two-parametric analysis of Rice-distributed data by means of the maximum likelihood approach // Am. J. Theor. Appl. Stat., 2013. Vol. 2. No. 3. P. 67–79.
17. Яковлева Т. В., Кульберг Н. С. Особенности функции правдоподобия статистического распределения Райса // Докл. РАН, 2014. Т. 457. № 4. С. 394–397.
18. Яковлева Т. В. Условия применимости статистической модели Райса и расчет параметров райсовского сигнала методом максимума правдоподобия // Компьютерные исследования и моделирование, 2014. Т. 6. № 1. С. 13–25.
19. Кульберг Н. С., Яковлева Т. В. Расчет параметров райсовского сигнала методами математической статистики. Свидетельство о государственной регистрации программы для ЭВМ № 2014616282 от 19.06.2014.

Поступила в редакцию 09.06.14

## TWO-PARAMETRIC ANALYSIS OF MAGNETIC-RESONANCE IMAGES BY THE MAXIMUM LIKELIHOOD TECHNIQUE IN COMPARISON WITH THE ONE-PARAMETRIC APPROXIMATION

*T. V. Yakovleva and N. S. Kulberg*

Dorodnitsyn Computing Center, Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The paper considers a method of the two-parametric analysis of magnetic-resonance image's data which allows getting a joint estimation of both the useful signal and the noise within the image being analyzed on the basis of the maximum likelihood principle. This technique presents an essentially new approach to data processing in the conditions of the Rice distribution and can be efficiently used in information technologies for solving a wide range of tasks connected with the Rician signals' filtering. Solving the two-parametric task is based on measured samples' data only and is not connected with *a priori* suppositions concerning the noise value which inevitably limit the precision of the one-parametric method. By means of computer simulation, a comparative analysis of the traditional one-parametric method and the elaborated two-parametric method is made for estimation of the useful component of the signal forming the magnetic-resonance image. The statistical data for the shift and the spread of the sought-for signal and noise parameters are calculated. The advantages of the two-parametric technique of data analysis are shown.

**Keywords:** Rice distribution; likelihood function; maximum likelihood method; noise dispersion; signal-to-noise ratio

**DOI:** 10.14357/08696527140307

### References

1. Nakagami, M. 1940. Study of resultant amplitude of many vibrations whose phases and amplitudes are at random. *Nippon Elec. Commun. Eng.* 22:69–92.
2. Rice, S. O. 1944. Mathematical analysis of random noise. *Bell Syst. Tech. J.* 23:282.
3. Henkelman, R. M. 1985. Measurement of signal intensities in the presence of noise in MR images. *Med. Phys.* 12(2):232–233.
4. Wang, T., and T. Lei. 1994. Statistical analysis of MR imaging and its application in image modeling. *IEEE Conference (International) Image Processing and Neural Networks Proceedings.* I:866–870.
5. Gudbjartsson, H., and S. Patz. 1995. The Rician distribution of noisy MRI data. *Magnet. Reson. Med.* 34:910–914.
6. Sijbers, J., A. J. den Dekker, P. Scheunders, and D. Van Dyck. 1998. Maximum-likelihood estimation of Rician distribution parameters. *IEEE Trans. Med. Imaging* 17(3):357–361.

7. Carobbi, C. F. M. and M. Cati. 2008. The absolute maximum of the likelihood function of the Rice distribution: Existence and uniqueness. *IEEE Trans. Instrum. Meas.* 57(4):682–689
8. Sheil, W. C. 2012. Magnetic resonance imaging (MRI scan). Available at: [MedicineNet.com](#). Retrieved April 27, 2012.
9. Perov, A. I. 2003. *Statisticheskaya teoriya radiotekhnicheskikh system* [Statistical theory of radio-technical systems]. Moscow: Radiotekhnika. 400 p.
10. Marcum, J. I., and P. Swelling. 1960. Studies of target detection by pilled radar. *IRE Trans. Inform. IT-6*:59–267.
11. Gardner, L. A. 1966. Statistical methods for the analysis of reentry vehicle radar data. Lexington: M.I.T. Lincoln Lab. Part 1. Mass. Tech. Note 1966-5.
12. Talukdar, K. K., and W. D. Lawing. 1991. Estimation of the parameters of Rice distribution. *J. Acoust. Soc. Am.* 89(3):1193–1197.
13. Stuber, G. L. 2001. *Principles of mobile communication*. 2nd ed. Norwell, MA: Kluwer. 752 p.
14. Yakovleva, T. V., and A. V. Knyaz'kov. 2014. Dvukhparametricheskiy metod momen-tov obrabotki rezul'tatov periodicheskoy modulyatsii otrazhennogo sveta dlya otsenki EO koeffitsientov [Two-parametric method of moments for processing the results of the reflected light's periodical modulation at electrooptical coefficients' estimation]. Conference "Lasers Measurements Information" Proceedings. St. Petersburg. 39.
15. Benedict, T. R., and T. T. Soong. 1967. The joint estimation of signal and noise from the sum envelope. *IEEE Trans. Inform. Theory IT-13*(3):447–454.
16. Yakovleva, T. V., and N. S. Kulberg. 2013. Noise and signal estimation in MRI: Two-parametric analysis of Rice-distributed data by means of the maximum likelihood approach. *Am. J. Theor. Appl. Stat.* 2(3):67–79.
17. Yakovleva, T. V., and N. S. Kulberg. 2014. Osobennosti funktsii pravdopodobiya statisticheskogo raspredeleniya Raysa [The likelihood function's peculiarities for Rice statistical distribution]. *Dokl. RAS* 457(4):1–4.
18. Yakovleva, T. V. 2014. Usloviya primenimosti statisticheskoy modeli Raysa i raschet parametrov Raysovskogo signala metodom maksimuma pravdopodobiya [Conditions of Rice statistical model applicability and estimation of the Rician signal's parameters by maximum likelihood technique]. *Komp'yuternye Issledovaniya i Modelirovanie* [Computer Research and Simulation] 6(1):13–25.
19. Kulberg, N. S., and T. V. Yakovleva. 2014. Raschet parametrov Raysovskogo signala metodami matematicheskoy statistiki [Calculation of Rician signal's parameters by the methods of mathematical statistics]. Certificate on the state registration of the computer program No. 2014616282 of 19.06.2014, owner: CCAS.

Received June 09, 2014

## Contributors

**Yakovleva Tatiana V.** (b. 1956) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Dorodnicyn Computing Center, Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; [yakovleva@ccas.ru](mailto:yakovleva@ccas.ru)  
**Kulberg Nicholas S.** (b. 1970) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Dorodnicyn Computing Center, Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; [kulberg@yandex.ru](mailto:kulberg@yandex.ru)

## СВОЙСТВА ОКОННОЙ ДИСПЕРСИИ МИОГРАММЫ КАК СЛУЧАЙНОГО ПРОЦЕССА

*М. Ш. Хазиахметов<sup>1</sup>*

**Аннотация:** В настоящее время ведется значительный объем работ по исследованию активности головного мозга. Один из подходов в них — метод вызванных потенциалов: испытуемый многократно повторяет некоторые действия, а его мозговая активность и некоторые вспомогательные сигналы записываются и затем анализируются с целью выявления необходимых зависимостей. В ряде работ, выполненных автором совместно с коллегами, рассматриваются алгоритмы локализации опорных точек в миограмме (и некоторых других сигналах), соответствующих началам движения в конкретном эксперименте, в котором человек совершает движение пальцем руки. Точность определения этих точек критична для дальнейшей обработки магнитоэнцефалограммы (МЭГ). В данной статье рассматриваются теоретические аспекты данных алгоритмов, а именно свойства миограммы и ее оконной дисперсии как случайных процессов.

**Ключевые слова:** выборочная оконная дисперсия; миограмма; магнитоэнцефалограмма

**DOI:** 10.14357/08696527140308

### 1 Введение

Приведем схему эксперимента, порядок работы с полученными в его ходе данными, а также краткое описание рассматриваемых алгоритмов нахождения опорных точек.

Испытуемый кладет ладонь на стол и в течение нескольких минут выполняет одно и то же движение — удар указательным пальцем по его поверхности. При этом производится запись мозговой активности (МЭГ-сигнал), мышечной деятельности (миограмма). Дополнительно также могут быть записаны ускорение пальца (сигнал акселерометра) и периоды касания стола (сигнал с кнопки на основе фотоэлемента).

Различить по МЭГ периоды движения и покоя не представляется возможным ввиду малого отношения сигнал / шум. Сделать это позволяют все перечисленные выше дополнительные сигналы. Однако ключевым при работе с МЭГ для нейрофизиологов является не столько сам факт движения, сколько момент его

---

<sup>1</sup>Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, khaziakhmetov@yandex.ru

начала. Под истинным моментом начала движения понимается момент, когда мозг отдает команду на его выполнение. Определить его точно невозможно. Вместе с тем, считая, что время распространения импульса от мозга до мышцы одинаково, можно в менее зашумленных дополнительных сигналах выделить опорные точки (моменты начала проявления вызванной активности) и работать с ними (константное смещение по времени не имеет значения в контексте работы с МЭГ). С точки зрения точности локализации опорных точек наибольшую ценность представляет миограмма: смещение по времени в ней относительно МЭГ можно считать постоянным (у других — нет) и оно наименьшее среди указанных сигналов; кроме того, показания с акселерометра и кнопки снимаются не всегда.

Обработка МЭГ заключается в усреднении ее отрезков вокруг опорных точек с целью выделения канала МЭГ с максимальным откликом — этот канал соответствует датчику, находящемуся ближе всего к отделу мозга, «запустившему» движение. По сути, это один из методов решения так называемой обратной задачи — определения местоположения источника сигнала по данным с множества принимающих устройств. Есть и другие подходы к решению этой задачи (см., например, [1, 2]).

Выделение опорных точек по сигналу с кнопки никаких проблем не вызывает. Все дальнейшее рассмотрение будет производиться для миограммы, так как обработка данных акселерометра производится точно так же.

Алгоритм локализации опорных точек, выбранный в качестве базового, сравнительно прост и основан на том факте, что значения миограммы в состоянии покоя и движения заметно отличаются. Вся миограмма рассматривается как случайный процесс  $M_t$ , для которого строится оконная дисперсия  $W_{n,t}$  ( $n$  — ширина окна), — это усиливает отличия в значениях и избавляет от трендов, причем резкий рост оконной дисперсии миограммы (ОДМ) будет начинаться именно в момент перехода процесса из одного состояния в другое. Далее строится эмпирическая функция распределения (ЭФР) ОДМ, по которой вычисляется квантиль достаточно высокого порядка. Все моменты пересечения отрезков возрастания ОДМ и квантили после очевидных фильтраций (необходимо исключить очень близкие точки) считаются опорными точками. Более подробное рассмотрение приведено в [3, 4].

Практика использования данного алгоритма показала, что он требователен к выбору квантили. Но даже сравнительно удачный выбор для конкретного испытуемого все равно может приводить к потере значительного числа опорных точек либо, наоборот, к большому числу ложных срабатываний. Это связано с изначальной неоднородностью данных — характер движений человека в ходе эксперимента меняется (хотя бы из-за того, что он постепенно устает), сколь-нибудь одинаковы только несколько последовательных движений (человека предупреждают о необходимости делать одинаковые движения, насколько это возможно). Указанные свойства были практически установлены в ходе выполнения работы [5].

Модифицированная версия алгоритма (подробное рассмотрение в [6]) при помощи сравнительно высокой квантили находит опорные точки базовым алгоритмом (этим достигается изначальная малая вероятность ложных срабатываний). Далее по ним переразбивает миограмму на сравнительно короткие отрезки (так, чтобы движения не попадали на их края), на которых вновь отдельно для каждого из них применяет базовый алгоритм с пересчитанной эмпирической функцией распределения и квантилью соответственно (это позволяет разбить сигнал на участки однородности, где можно даже с относительно низкой квантилью выявлять опорные точки с малой вероятностью ложного срабатывания). Все определенные таким образом точки считаются предварительными. Каждой точке соответствует предшествующий интервал покоя — на нем вновь вычисляется ЭФР ОДМ и достаточно высокая квантиль, и опорной точкой считается последний момент времени до предварительной точки, когда ОДМ меньше этой квантили. Последний шаг существенно повышает точность, что зафиксировано при вычислении отклика по усреднению МЭГ. Это связано, прежде всего, с тем, что (как продемонстрировано в [7]) базовый алгоритм помещает опорную точку в практически произвольную позицию как по времени относительно момента начала роста ОДМ, так и по амплитуде относительно локального максимума ОДМ.

## 2 Миограмма и ее оконная дисперсия как случайные процессы

Миограмма в оцифрованном варианте представляет собой временной ряд  $M_t$ , где  $t \in \{0, 1, \dots, T - 1\}$ . Условно ее можно разделить на чередующиеся между собой отрезки, соответствующие неподвижному состоянию пальца испытуемого и движению. Период времени, представляющий собой последовательную пару отрезков покоя и движения, будем называть эпохой. Границы эпох представляют собой множество точек вида

$$\{\tau_0, \tau_1, \dots, \tau_k\} \subset \mathbb{N} : 0 = \tau_0 < \tau_1 < \dots < \tau_k = T.$$

Тогда  $\{\tau_{j-1}, \tau_{j-1} + 1, \dots, \tau_j - 1\}$  — точки  $j$ -й эпохи ( $j = \overline{1, k}$ ).

Для каждой эпохи представим наблюдаемое значение миограммы  $M_t$  как сумму истинного неслучайного значения  $m_t$  и шума  $\xi_t$ , имеющего нормальное распределение с нулевым математическим ожиданием и дисперсией  $\sigma_j^2$ , причем  $\{\xi_{\tau_{j-1}}, \xi_{\tau_{j-1}+1}, \dots, \xi_{\tau_j-1}\}$  — совокупность независимых одинаково распределенных случайных величин (н. о. р. с. в.), а  $\{\xi_0, \xi_1, \dots, \xi_{T-1}\}$  — совокупность независимых случайных величин (н. с. в.):

$$\forall j = \overline{1, k} \begin{cases} \forall t \in \{\tau_{j-1}, \tau_{j-1} + 1, \dots, \tau_j - 1\} \begin{cases} M_t = m_t + \xi_t; \\ \xi_t \sim N(0, \sigma_j^2); \end{cases} \\ \{\xi_{\tau_{j-1}}, \xi_{\tau_{j-1}+1}, \dots, \xi_{\tau_j-1}\} \text{ — н. о. р. с. в.}; \\ \{\xi_0, \xi_1, \dots, \xi_{T-1}\} \text{ — н. с. в.} \end{cases} \quad (1)$$

Заметим, что в реальной миограмме нет жесткой границы между эпохами — в принципе, граничную точку можно поставить в любом месте на периоде покоя. С учетом особенностей работы модифицированной версии алгоритма граничные точки выбираются по возможности меньшими, чтобы обеспечить достаточный интервал значений ОДМ для вычисления квантили шума. Кроме того, нет и резких переходов от одного значения дисперсии шума миограммы к другому. Благодаря этому отрезки миограммы на стыке эпох можно считать статистически соответствующими с точки зрения распределения шума обеим эпохам, а потому, если в дальнейшем при расчете характеристик одной эпохи используются первые несколько точек из следующей, никак не учитывается тот факт, что формально с точки зрения предложенной в (1) модели они имеют другое распределение. Сформулируем это в следующем виде:

**Допущение.** Если  $n$  — ширина окна при расчете скользящего среднего и ОДМ миограммы, тогда будем считать, что для каждой эпохи, начиная со второй, первые  $n - 1$  точек можно использовать при вычислении значений скользящего среднего и ОДМ предыдущей эпохи с ее параметрами распределения шума.

Пусть  $n$  — ширина окна при расчете ОДМ. Тогда на каждой эпохе для соответствующего скользящего среднего  $\overline{M}_{n,t}$  наблюдаемого сигнала будем иметь представление в терминах скользящих средних истинного сигнала  $\overline{m}_{n,t}$  и шума  $\overline{\xi}_{n,t}$ :

$$\overline{M}_{n,t} = \frac{1}{n} \sum_{i=0}^{n-1} M_{t+i} = \frac{1}{n} \sum_{i=0}^{n-1} m_{t+i} + \frac{1}{n} \sum_{i=0}^{n-1} \xi_{t+i} = \overline{m}_{n,t} + \overline{\xi}_{n,t}. \quad (2)$$

Оконная дисперсия миограммы по определению имеет вид:

$$W_{n,t} = \frac{1}{n} \sum_{i=0}^{n-1} (M_{t+i} - \overline{M}_{n,t})^2. \quad (3)$$

**Лемма.** Для ОДМ, соответствующей модели (1) со сделанным выше допущением, справедливо разложение:

$$W_{n,t} = \frac{1}{n} \sum_{i=0}^{n-1} m_{t+i}^2 - \overline{m}_{n,t}^2 + \frac{1}{n} \sum_{i=0}^{n-1} \xi_{t+i}^2 - \overline{\xi}_{n,t}^2 + \frac{2}{n} \sum_{i=0}^{n-1} \xi_{t+i} (m_{t+i} - \overline{m}_{n,t}). \quad (4)$$

**Доказательство.** Подстановкой (1) и (2) в (3) получаем:

$$\begin{aligned} W_{n,t} &= \frac{1}{n} \sum_{i=0}^{n-1} ((m_{t+i} - \overline{m}_{n,t}) + (\xi_{t+i} - \overline{\xi}_{n,t}))^2 = \frac{1}{n} \sum_{i=0}^{n-1} (m_{t+i} - \overline{m}_{n,t})^2 + \\ &\quad + \frac{1}{n} \sum_{i=0}^{n-1} (\xi_{t+i} - \overline{\xi}_{n,t})^2 + \frac{2}{n} \sum_{i=0}^{n-1} (m_{t+i} - \overline{m}_{n,t}) (\xi_{t+i} - \overline{\xi}_{n,t}). \end{aligned} \quad (5)$$

Таким образом, ОДМ распадается на три компонента:

- (1)  $A_{n,t} = (1/n) \sum_{i=0}^{n-1} (m_{t+i} - \overline{m}_{n,t})^2$  — описывает неслучайную часть ОДМ (фактически это оконная дисперсия сигнала при отсутствии шума);
- (2)  $B_{n,t} = (1/n) \sum_{i=0}^{n-1} (\xi_{t+i} - \overline{\xi}_{n,t})^2$  — описывает компонент ОДМ, возникший исключительно за счет шума в миограмме (оконная дисперсия шума);
- (3)  $C_{n,t} = (2/n) \sum_{i=0}^{n-1} (m_{t+i} - \overline{m}_{n,t}) (\xi_{t+i} - \overline{\xi}_{n,t})$  — описывает компонент ОДМ, возникающий как суперпозиция истинного сигнала и шума.

Рассмотрим 1-й компонент:

$$\begin{aligned} A_{n,t} &= \frac{1}{n} \left( \sum_{i=0}^{n-1} m_{t+i}^2 + \sum_{i=0}^{n-1} \overline{m}_{n,t}^2 - 2 \sum_{i=0}^{n-1} m_{t+i} \overline{m}_{n,t} \right) = \\ &= \frac{1}{n} \left( \sum_{i=0}^{n-1} m_{t+i}^2 + n \overline{m}_{n,t}^2 - 2n \overline{m}_{n,t}^2 \right) = \\ &= \frac{1}{n} \left( \sum_{i=0}^{n-1} m_{t+i}^2 - n \overline{m}_{n,t}^2 \right) = \frac{1}{n} \sum_{i=0}^{n-1} m_{t+i}^2 - \overline{m}_{n,t}^2. \quad (6) \end{aligned}$$

Для 2-го компонента по аналогии с (6) можем записать:

$$B_{n,t} = \frac{1}{n} \sum_{i=0}^{n-1} \xi_{t+i}^2 - \overline{\xi}_{n,t}^2. \quad (7)$$

Рассмотрим 3-й компонент:

$$\begin{aligned} C_{n,t} &= \frac{2}{n} \sum_{i=0}^{n-1} \xi_{t+i} (m_{t+i} - \overline{m}_{n,t}) - \frac{2}{n} \overline{\xi}_{n,t} \sum_{i=0}^{n-1} (m_{t+i} - \overline{m}_{n,t}) = \\ &= \frac{2}{n} \sum_{i=0}^{n-1} \xi_{t+i} (m_{t+i} - \overline{m}_{n,t}). \quad (8) \end{aligned}$$

Объединяя (6)–(8), получаем (4).  $\square$

**Определение.** Назовем шумовой составляющей ОДМ, соответствующей модели (1) со сделанным выше допущением, сумму компонентов  $B_{n,t}$  и  $C_{n,t}$ , полученных в лемме.

**Теорема.** В точке  $t$ , принадлежащей  $j$ -й эпохе ( $j = \overline{1, k}$ ), шумовая составляющая ОДМ, соответствующей модели (1) со сделанным выше допущением, представляет собой либо случайную величину с распределением  $\Gamma\left(n/(2\sigma_j^2), (n-1)/2\right)$ , либо сумму двух случайных величин с распределениями  $\Gamma\left(n/(2\sigma_j^2), (n-1)/2\right)$  и  $N\left(0, (4\sigma_j^2/n^2) \sum_{i=0}^{n-1} (m_{t+i} - \bar{m}_{n,t})^2\right)$ .

**Доказательство.** Исходя из того, что 2-й компонент разложения ОДМ в лемме — оконная дисперсия нормально распределенной случайной величины, можем записать

$$\frac{nB_{n,t}}{\sigma_j^2} \sim \chi_{n-1}^2 = \Gamma\left(\frac{1}{2}, \frac{n-1}{2}\right)$$

для  $t \in \{\tau_{j-1}, \tau_{j-1} + 1, \dots, \tau_j - 1\}$ . Тогда получаем, что

$$B_{n,t} \sim \Gamma\left(\frac{n}{2\sigma_j^2}, \frac{n-1}{2}\right).$$

Используя (1), а также свойство усиленной воспроизводимости нормального распределения, получаем распределение 3-го компонента разложения ОДМ в лемме при условии, что  $\exists i, i = \overline{0, n-1} : m_{t+i} - \bar{m}_{n,t} \neq 0$ , т. е. когда среди истинных значений миограммы в точках  $t, t+1, \dots, t+n-1$  есть различные:

$$\begin{aligned} C_{n,t} &\sim N\left(0, \frac{4\sigma_j^2}{n^2} \sum_{i=0}^{n-1} (m_{t+i} - \bar{m}_{n,t})^2\right) = \\ &= N\left(0, \frac{4\sigma_j^2}{n} A_{n,t}\right), \quad \forall t \in \{\tau_{j-1}, \tau_{j-1} + 1, \dots, \tau_j - 1\}. \end{aligned}$$

Таким образом, дисперсия  $C_{n,t}$  линейно зависит от  $A_{n,t}$  — оконной дисперсии «чистого» сигнала.

В случае, когда истинный сигнал в точках  $\tau_{j-1}, \tau_{j-1} + 1, \dots, \tau_j - 1$  неизменен, получаем, что  $C_{n,t} = 0$ , а шумовая составляющая ОДМ состоит только из гамма-распределенного компонента  $B_{n,t}$ . В противном случае в разложении (4) присутствуют оба компонента.  $\square$

Резюмируя полученные в (5)–(8) соотношения, получаем итоговый результат:

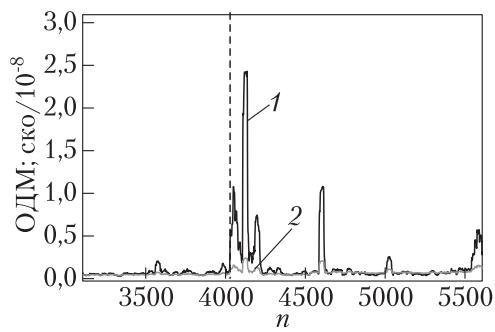
$$\forall j \in \overline{1, k} \quad \forall t \in \{\tau_{j-1}, \tau_{j-1} + 1, \dots, \tau_j - 1\} \left\{ \begin{array}{l} W_{n,t} = A_{n,t} + B_{n,t} + C_{n,t}; \\ A_{n,t} = \frac{1}{n} \sum_{i=0}^{n-1} m_{t+i}^2 - \overline{m_{n,t}}^2; \\ B_{n,t} = \frac{1}{n} \sum_{i=0}^{n-1} \xi_{t+i}^2 - \overline{\xi_{n,t}}^2; \\ C_{n,t} = \frac{2}{n} \sum_{i=0}^{n-1} \xi_{t+i} (m_{t+i} - \overline{m_{n,t}}); \\ B_{n,t} \sim \Gamma \left( \frac{n}{2\sigma_j^2}, \frac{n-1}{2} \right); \\ C_{n,t} \sim N \left( 0, \frac{4\sigma_j^2}{n} A_{n,t} \right) \text{ при } A_{n,t} > 0, \\ C_{n,t} = 0 \text{ иначе.} \end{array} \right.$$

Полученный набор соотношений демонстрирует, что

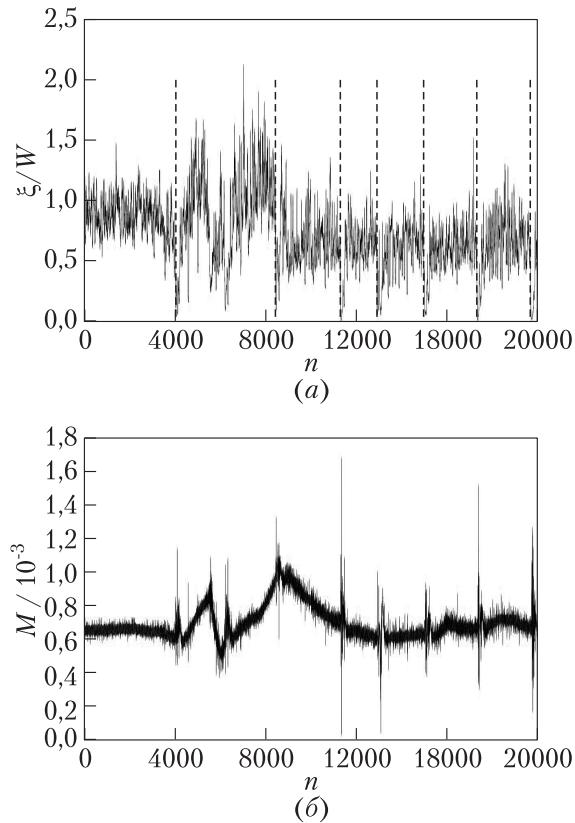
- (1) в наблюдаемых значениях оконной дисперсии присутствует шумовой компонент  $B_{n,t}$  — случайная величина, распределение которой не зависит от истинных значений миограммы, однако, поскольку шум на каждой эпохе (последовательной паре покой–движение) имеет свою дисперсию, зависимость от времени все равно есть;
- (2) шумовой компонент  $C_{n,t}$  имеет нормальное распределение, дисперсия которого зависит от истинных значений миограммы и дисперсии шума на эпохе. Данный компонент может отсутствовать в разложении, если «истинная» миограмма постоянна в пределах окна  $t, t+n-1$ ;
- (3) так как дисперсии  $B_{n,t}$  и  $C_{n,t}$  (если присутствует в разложении) равны соответственно  $(2(n-1)/n^2)\sigma_j^4 (\approx 2\sigma_j^4/n)$  и  $(4\sigma_j^2/n)A_{n,t}$ , то преобладание одного из компонентов над другим определяется взаимным соотношением  $\sigma_j^2$  и  $A_{n,t}$ . Учитывая природу исходного сигнала, можно утверждать, что в состоянии покоя влияние этих компонентов сравнимо либо преобладает  $B_{n,t}$ , а во время совершения движения преобладает  $C_{n,t}$ .

### 3 Заключение

Проиллюстрируем полученные утверждения на основе данных одной из миограмм. Для этого по точкам, расставленным модифицированным алгоритмом,



**Рис. 1** Окночная дисперсия миограммы (1) и оценка скна шума в ОДМ (2)



**Рис. 2** Отношение оценки шума в ОДМ к ОДМ (а) и миограмма (б)

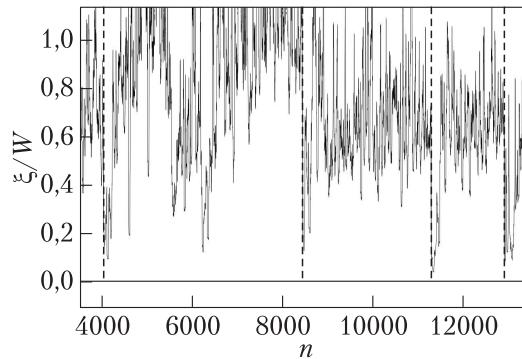


Рис. 3 Отношение оценки шума в ОДМ к ОДМ

определим периоды покоя, соответствующие движениям, и по ним оценим для каждой эпохи дисперсию шума миограммы. Далее вычислим значения дисперсии  $B_{n,t}$  — это кусочно-постоянная функция с точками разрыва на границах эпох. Взяв в качестве оценки  $A_{n,t}$  ОДМ, получим оценку для дисперсии  $C_{n,t}$  (шум, главным образом, увеличивает ОДМ, поэтому в каком-то смысле это оценка сверху). Зная, что дисперсия суммы оценивается удвоенной суммой дисперсий, получаем оценку для среднего квадратичного отклонения (ско) суммы шумовых компонент. На рис. 1 показана ОДМ с шириной окна 30 и оценка скро шума в ней, вертикальная штриховая черта — опорная точка. Видно, что в состоянии покоя характерные отклонения шума сопоставимы с ОДМ, в то время как во время движения энергия сигнала выше.

На рис. 2 и 3 показано отношение оценки скро шума в ОДМ к самой ОДМ. В первом случае дополнительно приведена миограмма, во втором — верхний график значительно увеличен. Опорные точки отмечены вертикальными штриховыми чертами. Нетрудно заметить, что в момент начала движения оценка характерного отклонения, получаемого за счет шума, не превышает 20% наблюдаемого значения ОДМ. Этот факт, по сути, гарантирует, что в правой полуокрестности опорной точки ОДМ будет обладать значительной тенденцией к возрастанию, что и используют оба алгоритма.

## Литература

1. Захарова Т. В., Гончаренко М. Б., Никифоров С. Ю. Метод решения обратной задачи магнитоэнцефалографии, основанный на кластеризации поверхности мозга // Статистические методы оценивания и проверки гипотез: Межвузовский сб. науч. тр. — Пермь: ПГНИУ, 2013. Т. 25. С. 120–125.
2. Бенинг В. Е., Драницына М. А., Захарова Т. В., Карпов П. И. Решение обратной задачи в многодипольной модели источников магнитоэнцефалограмм методом независимых компонент // Информатика и её применения, 2014. Т. 8. Вып. 2. С. 79–87.

3. *McGillem C. D., Aunon J. I.* Analysis of event-related potentials // Methods of analysis of brain electrical and magnetic signals: EEG handbook / Eds. A. Gevins, A. Rémond. — Amsterdam: Elsevier Science Publs., 1987. P. 131–169.
4. *Fabiani M., Gratton G., Federmeier K. D.* Event-related brain potentials: Methods, theory and applications // Handbook of psychophysiology / Eds. J. T. Cacioppo, L. G. Tassinary, G. G. Berntson. — Cambridge: Cambridge University Press, 2007. P. 85–119.
5. *Dranitsyna M., Zakharova T., Allakhverdiyeva V., Chshenyavskaya E.* Probability density function of myogram noise and its role in localization of brain activity sources // 32nd Seminar (International) on Stability Problems for Stochastic Models (ISSPSM'2014) and 8th Workshop (International) “Applied Problems in Theory of Probabilities and Mathematical Statistics Related to Modeling of Information Systems” (APTP + MS'2014): Book of abstracts. — Moscow: IPI RAN, 2014. P. 25–26.
6. Захарова Т. В., Никифоров С. Ю., Гончаренко М. Б., Драницына М. А., Климов Г. А., Хазиахметов М. Ш., Чаянов Н. В. Методы обработки сигналов для локализации невосполнимых областей головного мозга // Системы и средства информатики, 2012. Т. 22. № 2. С. 157–175.
7. Хазиахметов М. Ш., Захарова Т. В. Об алгоритмах нахождения опорных точек миограммы для использования в локализации невосполнимых областей головного мозга // Статистические методы оценивания и проверки гипотез: Межвузовский сб. науч. тр. — Пермь: ПГНИУ, 2013. Т. 25. С. 56–63.

Поступила в редакцию 20.08.14

---

## PROPERTIES OF WINDOW DISPERSION OF MYOGRAM AS A STOCHASTIC PROCESS

*M. Sh. Khaziakhmetov*

Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation

**Abstract:** Nowadays, the research of brain activity is conducted by many scientists. One of the approaches is the method of caused potentials: the testee repeats some actions, his brain activity and some additional signals are registered and then are analysed in order to detect necessary dependences. A number of works of the author and his colleagues consider algorithms of localization of reference points in a myogram (and some other signals) corresponding to the movement beginnings in the experiment in which a person makes movements by a hand finger. Accuracy of localization of these points is critical for further processing of magnetoencephalogram. The paper covers some theoretical aspects of these algorithms such as models of a myogram and its window dispersion as stochastic processes.

**Keywords:** window dispersion; myogram; magnetoencephalogram

**DOI:** 10.14357/08696527140308

## References

1. Zakharova, T. V., M. B. Goncharenko, and S. Yu. Nikiforov. 2013. Metod resheniya obratnoy zadachi magnitoentsefalografi, osnovannyy na klasterizatsii poverkhnosti mozga [Method for solving the inverse problem of magnetoencephalography, based on brain surface clustering]. *Statisticheskie metody otsenivaniya i proverki gipotez: Mezhevuzovskiy sbornik nauchnykh trudov* [Statistical methods of estimation and hypothesis testing: Interuniversity collection of scientific papers]. Perm: Perm State National Research University. 25:120–125.
2. Bening, V. E., M. A. Dranitsyna, T. V. Zakharova, and P. I. Karpov. 2014. Reshenie obratnoy zadachi v mnogodipol'noy modeli istochnikov magnitoentsefalogramm metodom nezavisimykh komponent [Independent component analysis for inverse problem in multidipole model magnetoencephalogram's sources]. *Informatika i ee Primeneniya—Inform. Appl.* 8(2):79–87. doi: 10.14357/19922264140208.
3. McGillem, C. D., and J. I. Aunon. 1987. Analysis of event-related potentials. *Methods for analysis of brain electrical and magnetic signals. EEG Handbook* (revised series, 1). Eds. A. Gevins and A. Rémond. Amsterdam: Elsevier. 131–169.
4. Fabiani, M., G. Gratton, and K. D. Federmeier. 2007. Event-related brain potentials: Methods, theory and applications. *Handbook of psychophysiology*. Eds. J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson. Cambridge: Cambridge University Press. 85–119.
5. Dranitsyna, M., T. Zakharova, V. Allakhverdiyeva, and E. Chshenyavskaya. 2014. Probability density function of myogram noise and its role in localization of brain activity sources. *32nd Seminar (International) on Stability Problems for Stochastic Models (ISSPSM'2014) and 8th Workshop (International) “Applied Problems in Theory of Probabilities and Mathematical Statistics Related to Modeling of Information Systems” (APTP + MS'2014). Book of abstracts*. Moscow: IPI RAN. 25–26.
6. Zakharova, T. V., S. Yu. Nikiforov, M. B. Goncharenko, M. A. Dranitsyna, G. A. Klimov, M. Sh. Khaziakhmetov, and N. V. Chayanov. 2012. Metody obrabotki signalov dlya lokalizatsii nevopolnomykh oblastey golovnogo mozga [Signal processing methods for localization of nonrenewable brain regions]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 22(2):157–175.
7. Khaziakhmetov, M. Sh., and T. V. Zakharova. 2013. Ob algoritmakh nakhodeniya opornykh tochek miogrammy dlya ispol'zovaniya v lokalizatsii nevopolnomykh oblastey golovnogo mozga [About algorithms of localization of reference points in myogram for usage in nonrenewable brain regions detection]. *Statisticheskie metody otsenivaniya i proverki gipotez: Mezhevuzovskiy sbornik nauchnykh trudov* [Statistical methods of estimation and hypothesis testing: Interuniversity collection of scientific papers]. Perm: Perm State National Research University. 25:56–63.

*Received August 20, 2014*

## Contributor

**Khaziakhmetov Maxim Sh.** (b. 1989) — PhD student, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; khaziakhmetov@yandex.ru

## ИССЛЕДОВАНИЕ ЛАБОРАТОРНОГО ПРОТОТИПА ИСКУССТВЕННОЙ ГЕТЕРОГЕННОЙ СИСТЕМЫ ДЛЯ ДИАГНОСТИКИ АРТЕРИАЛЬНОЙ ГИПЕРТЕНЗИИ

*И. А. Кириков<sup>1</sup>, А. В. Колесников<sup>2</sup>, С. Б. Румовская<sup>3</sup>*

**Аннотация:** Представлены результаты имитации лабораторного прототипа функциональной гибридной интеллектуальной системы (ФГиИС) — Виртуального медицинского консилиума для диагностики артериальной гипертензии (АГ) — средствами пакета MATLAB 7 — Simulink. Приведены данные модельных экспериментов и их интерпретация. Описаны результаты тестирования функциональных моделей гетерогенного модельного поля распознавания электрокардиограмм (ЭКГ), суточного мониторирования артериального давления (АД) и модулей классификации состояния здоровья пациента, а также технологических модулей оценки полезности симптоматической информации для постановки диагноза генетическими алгоритмами (ГА). Эксперименты с лабораторным прототипом показывают возможность снижения порога «замаскированных и недовыявленных» диагнозов симптоматической АГ и эссенциальной АГ с 70% до 20% в результате применения Виртуального консилиума на практике в лечебно-профилактических учреждениях широкого профиля.

**Ключевые слова:** лабораторный прототип Виртуального медицинского консилиума; артериальная гипертензия; нейронные сети; генетические алгоритмы; нечеткие системы; функциональная гибридная интеллектуальная система

**DOI:** 10.14357/08696527140309

### 1 Введение

Артериальная гипертензия — основной сердечно-сосудистый фактор риска. Из проблем диагностики АГ, обсуждавшихся в работах [1, 2], для настоящего исследования актуальны две:

- (1) низкое качество диагностического процесса: диагноз АГ «замаскирован и недовыявлен» у 70% пациентов амбулаторно-клинического учреждения;
- (2) отставание принципов, закономерностей, методов и методик, выработанных теоретической и клинической медициной, медицинской информатикой, от

<sup>1</sup>Калининградский филиал Института проблем информатики Российской академии наук, baltbipiran@mail.ru

<sup>2</sup>Балтийский федеральный университет им. И. Канта; Калининградский филиал Института проблем информатики Российской академии наук, avkolesnikov@yandex.ru

<sup>3</sup>Калининградский филиал Института проблем информатики Российской академии наук, sophiyabr@gmail.com

современных картин мира и организма человека: врачи и системы поддержки принятия диагностических решений «видят» пациента односторонне, с позиций одной профессии, одной дисциплины, одной школы, что характерно для гомогенных систем и ведет к решению задачи диагностики АГ несистемно, когда разнообразные знания не взаимодействуют и не сотрудничают для достижения общей цели.

Последнее делает в информатике актуальной разработку искусственных гетерогенных диагностических систем со свойствами дополнительности, сотрудничества и относительности знаний, синтезирующих интегрированные методы и модели, разнообразие которых релевантно разнообразию показателей диагностики организма человека.

Для создания компьютерной системы поддержки принятия решений сложной задачи диагностики АГ применен метод ФГиИС [3] — искусственных гетерогенных систем, моделирующих коллективные решения (консилиумы), разнообразие хранимой и перерабатываемой информации в которых релевантно разнообразию информации в диагностических ситуациях. Такие системы, Виртуальные консилиумы, способны динамично синтезировать над разнообразием модельного поля новый метод выработки диагноза АГ для каждого пациента индивидуально.

Настоящая работа посвящена тестированию моделей из гетерогенного модельного поля и экспериментам с синтезируемым над ним лабораторным прототипом гибридной интеллектуальной системы средствами пакета MATLAB 7 — Simulink, а также интерпретации их результатов.

## **2 Тестирование гетерогенного модельного поля Виртуального консилиума**

Гетерогенное модельное поле — основополагающая составляющая ФГиИС [3]. Оно функционально избыточно, содержит модели, релевантные подзадачам из декомпозиции сложной диагностической задачи [4], и имитирует многообразие и относительность профессиональных экспертных знаний, имеющихся в распоряжении лица, принимающего решение (ЛПР), для организации консилиума. С учетом такого разнообразия ЛПР делает выбор о включении экспертов в состав консилиума и обеспечивает дополнительность и сотрудничество знаний в ходе коллективных рассуждений над диагностической задачей. Гетерогенное модельное поле Виртуального консилиума состоит из 21 модели, девять из которых — технологические элементы и используются для организации эффективной обработки данных и знаний [5].

### **2.1 Обучение и тестирование функционального элемента — искусственной нейронной сети для распознавания электрокардиограмм**

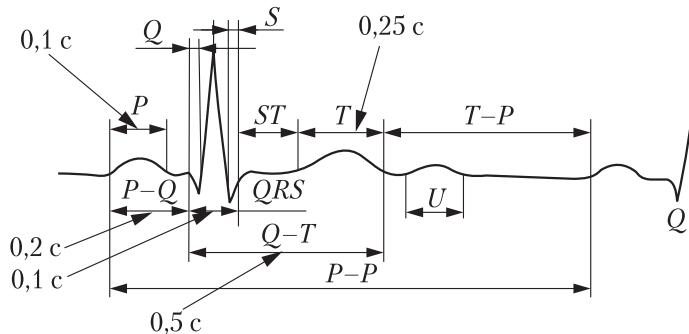
В [5] для искусственной нейронной сети (ИНС), «читающей» ЭКГ (модель ИНСРЭКГ), выбраны:

- (1) модульная структура блоков распознавания патологии — сигнал образа ЭКГ поступает на все модули одновременно;
- (2) топология — сеть без обратных связей;
- (3) архитектура — многослойная сеть с прямым распространением сигнала;
- (4) алгоритм обучения — «обратное распространение ошибки»;
- (5) парадигма обучения — обучение с учителем;
- (6) функции активации — логистическая функция;
- (7) тип нормализации — линейная;
- (8) число нейронов входного слоя — 128;
- (9) число нейронов в скрытом слое каждого из модулей изменяется от 1 до 51 с шагом 10 и выбирается по значению ошибки обучения каждого модуля, а также показателей чувствительности и специфичности, дающих клиническую и диагностическую интерпретацию решений, основанных на статистических характеристиках или логических правилах.

Чтобы свести до минимума сложность аппаратной части и процедуры регистрации, патологии распознаются по одноканальной записи *PQRST*-комплекса ЭКГ (рис. 1) в одном из 12 стандартных отведений [6]. Тем самым снижается стоимость регистрирующих комплексов ЭКГ и повышается эффективность их клинического применения.

Прохождение импульса по проводящей системе сердца графически записывается по вертикали в виде пиков — подъемов и спадов кривой линии. Эти пики называются зубцами и обозначаются латинскими буквами *P*, *Q*, *R*, *S* и *T*. На ЭКГ по горизонтали записывается время, в течение которого импульс проходит по определенным отделам сердца.

Для обучения с учителем и нейровычислений на вход ИНС подаются обучающая и тестирующая последовательности. Алгоритм создания базы данных



**Рис. 1** *PQRST*-комплекс ЭКГ

образов патологий сердца использует результаты [7, 8]. На вход алгоритма подаются файлы из архива PhysioBank — структурированного массива цифровых записей ЭКГ Национального института метрологии Германии (Physikalisch-Technische Bundesanstalt) [9]. Файлы сгруппированы в папки по патологиям и обрабатываются следующим образом:

- (1) запускается цикл с распознаванием  $R$ -зубцов в файле и сохранением их координат;
- (2) определяются  $RR$ -интервалы и изменяется частота дискретизации до 150 Гц; временное окно для анализа электрокардиосигнала (ЭКС) при извлечении  $RR$ -интервалов — 0,85 с, так как длительность левой и правой частей нормальной ЭКГ до и после появления зубца  $R$  равны соответственно 0,35 и 0,5 с; длительность зубца  $P \approx 0,1$  с, интервала  $PQ \approx 0,2$  с, интервала  $QS \approx 0,1$  с, зубца  $T \approx 0,25$  с [6]; частота 150 Гц выбрана как типичная для современных суточных мониторов [10];
- (3) определение опорных точек: границ зубцов  $P$ ,  $Q$ ,  $S$  и  $T$  [7], — и линейная нормализация компонентов образа ЭКС до диапазона нейронов от 0 до 1;
- (4) сохранение образов выделенных из файла кардиоциклов в файл с именем, содержащим номер патологии и метку назначения: обучение или тестирование.

В итоге формируется 14 файлов с обучающими и тестирующими последовательностями из примеров  $(X_1, \dots, X_{128}; Y_i)$ , где  $X_1, \dots, X_{128}$  — вектор образов ЭКС, компоненты которого — числа из интервала  $[0, 1]$ , построенный по записям кардиоциклов шести патологий и нормального состояния (всего выделено 400 кардиоциклов);  $Y_i$ ,  $i = \overline{1, 6}$ , — выход, соответствующий норме или одной из шести диагностируемых на ЭКГ патологий: принимает значение 1, если норма или патология имеют место, и 0 в противном случае. Результаты работы алгоритма сведены в табл. 1.

**Таблица 1** Состав базы данных патологий с разделением на обучающую и тестовую последовательности

Номер патологии	Тип патологии	Число обучающих примеров	Число тестирующих примеров
0	Норма	23	20
1	Желудочковые экстрасистолы	29	27
2	Отклонение электрической оси сердца влево	27	24
3	Блокада левой ножки пучка Гиса	35	31
4	Рубцовое поражение миокарда	37	34
5	Признаки перегрузки левого предсердия	30	27
6	Наджелудочковые нарушения ритма	31	28

Искусственная нейронная сеть тестируется по чувствительности, специфичности и ошибке обучения [8, 11]. Чувствительность ( $Sen$ ) вычисляется как отношение  $\langle$ числа истинно положительных решений $\rangle$  к  $\langle$ числу субъектов с заболеванием $\rangle$  и отражает способность ИНС обнаруживать заболевание. Положительное решение — ситуация положительного теста для пациента, имеющего заболевание. Специфичность ( $Sp$ ) вычисляется как отношение  $\langle$ числа истинно отрицательных решений $\rangle$  к  $\langle$ числу субъектов без заболевания $\rangle$  и отражает точность определения отсутствия заболевания. Отрицательное решение — ситуация отрицательного теста для пациента, не имеющего заболевания.

Ошибка обучения ИНС вычисляется алгоритмом обратного распространения ошибки по формуле [12]:

$$\delta_i^n = y_i^n - d_i, \quad (1)$$

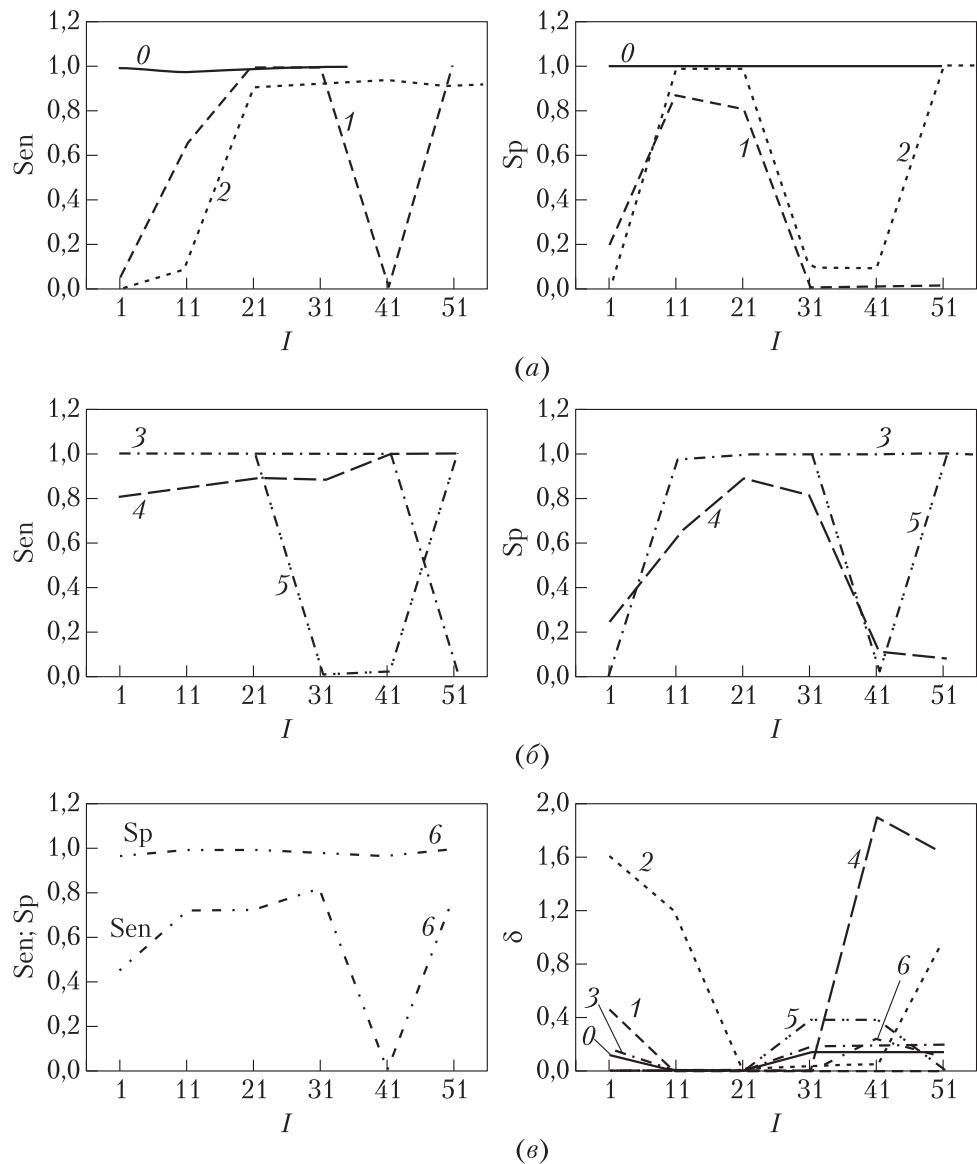
где  $y_i^n$  — значение выхода  $i$ -го нейрона выходного  $n$ -го слоя;  $d_i$  — требуемое значение  $i$ -го выхода,  $i = \overline{1, 6}$ .

Обучение и тестирование ИНС с оценкой  $Sen$ ,  $Sp$  и  $\delta$  в зависимости от числа нейронов в скрытых слоях каждого модуля реализовано поэтапно: (1) запуск цикла по перебору числа нейронов в скрытом слое каждого модуля нейросети с шагом 10 от 1 до 51; (2) нейровычисления с заданным числом нейронов скрытых слоев и вышеопределенными параметрами: тип нейросети, диапазоны входа, функции обучения и ошибки обучения, число слоев и нейронов в слое, функция активации слоев; (3) перебор и обучение модулей нейросети, затем тестирование и вычисление ошибки обучения  $Sen$  и специфичности  $Sp$ ; (4) исследование нейросети завершается сохранением показателей оценки работы ИНС для каждого заданного числа нейронов скрытого слоя: 1, 11, 21, 31, 41 и 51.

Результаты экспериментов с нейронной сетью для распознавания ЭКГ представлены на рис. 2.

Зависимости чувствительности  $Sen$  и специфичности  $Sp$  от числа нейронов скрытого слоя ( $I$ ) для модулей патологий под номерами 0–2, 3–5 и 6 (см. табл. 1) представлены на рис. 2, *a*, 2, *b* и 2, *v* соответственно. Зависимости ошибки обучения  $\delta$  от числа нейронов скрытого слоя для всех семи модулей — на рис. 2, *z*. Так как показатели чувствительности и специфичности — это меры достоверного обнаружения патологий и уверенности в их отсутствии, в идеале они должны быть близки к единице, а разность между ними стремиться к нулю. Поэтому значения для числа нейронов в скрытых слоях модулей выбирались таким образом, чтобы разница между  $Sen$  и  $Sp$  была минимальной либо, если разница ( $Sen - Sp$ ) одинакова, то была бы минимальной ошибка обучения. Как результат обучения и тестирования число нейронов скрытого слоя, значения  $Sen$ ,  $Sp$  и ошибки обучения равны:

- для модуля 0 — 11, 0,87, 0,98, 0,0023;
- для модуля 1 — 11, 0,91, 0,998, 0,0021;
- для модуля 2 — 21, 1, 1, 0,00024;



**Рис. 2** Результаты экспериментального исследования ИНСРЭКГ: 0 — модуль, соответствующий норме; 1 — модуль, распознающий желудочковые экстрасистолы на ЭКГ; 2 — отклонение электрической оси сердца влево; 3 — блокада левой ножки пучка Гиса; 4 — рубцовое поражение миокарда; 5 — признаки перегрузки левого предсердия; 6 — наджелудочковые нарушения ритма

- для модуля 3 — 21, 1, 0,998, 0,0002;
- для модуля 4 — 21, 0,89, 0,9, 0,0071;
- для модуля 5 — 11, 1, 1, 0,00025;
- для модуля 6 — 31, 0,823, 0,98, 0,00025.

Таким образом, после обучения ИНСРЭКГ средние значения чувствительности и специфичности равны 0,926 и 0,979 соответственно. Иными словами, в 93% случаев модель распознает наличие патологии, в 98% — отсутствие патологии, что указывает на релевантность модели реальной диагностической ситуации и позволяет включить модель в гетерогенное модельное поле для нейровычислений.

## 2.2 Обучение и тестирование функционального элемента — искусственной нейронной сети для расчета нормальных показателей суточного мониторирования артериального давления

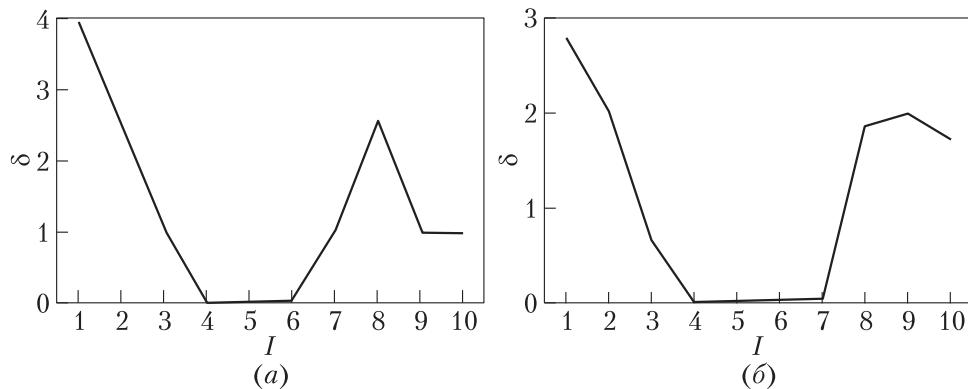
Модель ИНС для определения нормальных показателей суточного мониторирования АД (ИНССМАД) строит прогнозы, контролирует данные измерений суточного АД пациента и состоит из двух нейросетей: для систолического и для диастолического АД.

В [5] для этих нейросетей выбраны:

- топология — сети без обратных связей;
- архитектура — многослойные сети с прямым распространением сигнала, с обратным распространением ошибки и логистической функцией активации;
- число нейронов входного слоя — 2 (пол и возраст пациента) и выходного слоя — 24 нейрона, поскольку систолическое и диастолическое давление у пациента измеряется каждые полчаса с усреднением за час;
- исходное число нейронов в скрытом слое варьировалось в экспериментах от 1 до 10 и подбиралось по значению ошибки обучения (1) при  $i = 1, 24$ .

Для обучения с учителем на вход ИНС подавались обучающая и тестирующая последовательности из примеров  $(X_1, X_2; Y_{11}, \dots, Y_{124}; Y_{21}, \dots, Y_{224})$ , где  $X_1$  и  $X_2$  — это вектор входов: пол (0 — мужской, 1 — женский) и возраст (нормализуется до значений в интервале  $[0; 1]$ );  $Y_{11}, \dots, Y_{124}$  и  $Y_{21}, \dots, Y_{224}$  — выходы 1-й и 2-й нейросетей соответственно: систолическое и диастолическое давление (нормализуются до значений в интервале  $[0, 1]$ ). Обучающие и тестовые выборки сформированы из архива карт пациентов Калининградской областной клинической больницы (КОКБ) и представлены четырьмя файлами: два для обучения и два для тестирования. Для обучения и тестирования нейросеть моделируется с заданным числом нейронов скрытых слоев (1, 2, 3, 4, 5, 6, 7, 8, 9, 10), затем обучается и тестируется с вычислением ошибки обучения, сохраняемой вместе с заданным числом нейронов скрытого слоя.

Результаты экспериментов с ИНС для суточного мониторирования АД представлены на рис. З зависимостями средней ошибки обучения  $\delta$  от числа нейронов



**Рис. 3** Результаты экспериментального исследования ИНС мониторинга АД

скрытого слоя  $I$ : для модуля систолического давления по полу и возрасту больного — на рис. 3, а, а для модуля диастолического АД — на рис. 3, б.

Число нейронов в скрытых слоях обоих модулей выбиралось так, чтобы ошибка обучения была минимальной. В результате оптимальным оказалось число четыре с наименьшими средними ошибками обучения 0,021 и 0,015 для первого и второго модуля соответственно.

В итоге экспериментов среднее значение ошибки обучения для ИНССМАД составило 0,018, что позволило включить ее в гетерогенное модельное поле с оценкой 98% релевантности прогноза АД пациента диагностической ситуации. Такая оценка повышает вероятность достоверности заключительного диагноза, формируемого моделью, решающей подзадачу «Оценка степени и стадии АГ, степени риска сердечно-сосудистых заболеваний (ССЗ)» и использующим нейросетевой прогноз.

### 2.3 Исследование технологических элементов — генетических алгоритмов, решающих задачу построения информативного набора симптомов при диагностике заболеваний

Модели ГА подзадач «Построение информативного набора признаков (симптомов) при диагностике заболеваний» решают задачи безусловной оптимизации на множестве признаков, по которым диагностируются заболевания из декомпозиции сложной задачи диагностики АГ, и исключают из рассмотрения Виртуального консилиума избыточные и взаимозависимые признаки заболеваний. Это минимизирует общее число входных переменных (признаков, симптомов) на входе нечетких систем (НС), решающих задачи классификации патологического состояния здоровья пациента НСДАГ<sub>1</sub>–НСДАГ<sub>9</sub> (на рис. 4 (см. с. 134) этим подзадачам соответствуют модели  $\text{mod}_f^a \text{ДАГ}_1$  –  $\text{mod}_f^a \text{ДАГ}_9$ ).

У каждого классического ГА<sub>*k*</sub> [13] из ГАППС<sub>1–9</sub> размер индивидуума (*in<sub>k</sub>*) равен общему количеству учитываемых признаков в соответствующей подзадаче из ДАГ<sub>*k*</sub> для подзадач ППС<sub>*k*</sub> (*k* = 1, 9) — *in*<sub>1–9</sub> = 30, 20, 80, 20, 65, 65, 190, 130, 25 генам соответственно. Значение гена, равное 1, соответствует симптуому, принадлежащему множеству «Учитываемые признаки», и 0 — признаку, не включенному в это множество. Работает ГА<sub>*k*</sub> следующим образом:

- (1) инициализация — выбор начальной популяции индивидуумов в *N<sub>KB</sub>* = 20 особям;
- (2) каждый индивидуум накладывается как маска на *k*-ю НС, при этом в *k*-й НС учитываются только симптомы, принимающие значения 1 у индивидуума;
- (3) вычисляются функции приспособленности индивидуума как среднеквадратическая ошибка постановки диагноза по формуле:

$$f_i = \frac{1}{N_k} \sqrt{\sum_{l=1}^{N_k} \left( \left( \sum_{j=1}^N \delta_{lj} \right) / N \right)^2}, \quad (2)$$

где *i* =  $\overline{1, N_{KB}}$  — номер индивидуума в популяции; *j* — число наблюдений в тестовой выборке (*N* = 15); *f<sub>i</sub>* — среднеквадратическая ошибка классификации *k*-й НС на тестовой выборке из 15 наблюдений признаков состояния здоровья пациента;  $\delta_{ij}$  — ошибка классификации по *l*-му выходу *k*-й НС (равна 1, если совпадает с эталоном, и 0 — если нет); *N<sub>k</sub>* — число выходов *k*-й НС;

- (4) отбор родителей методом турнирной селекции: размер турнира — 2;
- (5) избранные порождают новые особи-потомки операторами мутации (вероятность 0,01) и скрещивания (вероятность 0,5), и формируется новая популяция;
- (6) после 50 итераций выбирается наилучшая особь с оптимальным набором признаков — это и есть маскирующая информация, аргумент функции модификации *k*-й НС (по умолчанию содержит весь набор признаков, отобранных экспертами).

Результаты работы ГА из ГАППС<sub>1–9</sub> исследовались в сравнении с данными табл. 2, содержащей объемы входной информации для НС, полученные в [14] по данным о коэффициентах достоверности СФ(*H, E*) отобранных экспертом признаков, вычислявшихся по теории подтверждения Карнапа [15].

Теория факторов уверенности основана на двух оценках: мере доверия МВ(*H, E*) и мере недоверия МД(*H, E*). Эти функции указывают степень увеличения доверия к гипотезе *H*, если факт *E* произошел, и степень увеличения недоверия к гипотезе *H*, если факт *E* имел место, соответственно. Значения

**Таблица 2** Оптимальный набор входных признаков для НСДАГ<sub>1–9</sub>

№	Наименование НС	Объем входной информации НС		
		Исходный	После обработки ГА	Вычисленный по коэффициенту достоверности
1	Диагностика поражений органов-мишней	30	25	26
2	Диагностика факторов риска АГ	20	20	20
3	Диагностика переброваскулярных болезней	80	64	68
4	Диагностика метаболического синдрома и сахарного диабета	20	14	15
5	Диагностика заболеваний периферических артерий	65	53	56
6	Диагностика ишемической болезни сердца	65	53	55
7	Диагностика эндокринной АГ	190	145	150
8	Диагностика паренхиматозной нефропатии	130	93	103
9	Диагностика реноваскулярной АГ	25	18	19

МВ(Н, Е) и МД(Н, Е) располагаются между нулем и единицей. Фактор уверенности CF(Н, Е) рассчитывается по формуле:

$$CF(H, E) = MB(H, E) - MD(H, E),$$

при этом CF(Н, Е) ∈ [−1, 1] и определяет степень доверия к гипотезе Н.

Анализ данных табл. 2 показывает, что объем входной информации для НС, вычисленный по коэффициенту достоверности, превышает объем, полученный после ГА. Иными словами, ГА исключили из рассмотрения большее количество симптомов, чем применение аппарата факторов уверенности. Анализ множеств симптомов обоими методами показал, что разница объясняется тем, что эксперт учитывает симптомы с достоверностью, равной нулю, не влияющие на постановку диагноза. В общей массе симптомов они так или иначе учитываются врачом. Генетические алгоритмы исключают из рассмотрения показатели здоровья с нулевым коэффициентом достоверности. Таким образом, ГА удовлетворительно минимизируют общее число входных переменных (симптомов) на входе НСДАГ<sub>1–НСДАГ<sub>9</sub></sub>, решая задачи классификации патологического состояния здоровья пациента. Это повышает эффективность Виртуального консилиума, поскольку уменьшается объем обрабатываемой информации о показателях состояния здоровья человека, вводимых в систему, а также снижает время опроса, осмотра и обследования пациента врачом.

## 2.4 Тестирование функциональных элементов гетерогенного модельного поля — нечетких систем, решающих подзадачи классификации патологического состояния здоровья пациента

Эти подзадачи сгруппированы в подзадачу «Диагностика критериев оценки сердечно-сосудистого риска и вторичной АГ у пациента» и подзадачу «Оценка степени и стадии артериальной гипертензии, степени риска ССЗ».

Модели НСДАГ<sub>1</sub>–НСДАГ<sub>9</sub> подзадач из группы «Диагностика критериев оценки сердечно-сосудистого риска и вторичной АГ у пациента» (ДАГ<sub>1–9</sub>) классифицируют патологическое состояния здоровья пациента.

Модель НССС подзадачи «Оценка степени и стадии артериальной гипертензии, степени риска ССЗ» (ССС) в соответствии с декомпозицией сложной задачи диагностики АГ [4] применяется последней в Виртуальном консилиуме и классифицирует:

- (1) степень АГ по данным систолического, диастолического давления и данным от ИНССМАД;
- (2) стадию эссенциальной АГ (или гипертонической болезни) в случае отсутствия вторичной АГ (диагностируется НСДАГ<sub>7–9</sub>) по данным от НСДАГ<sub>1–3</sub>, НСДАГ<sub>5</sub> и НСДАГ<sub>6</sub>;
- (3) общий риск ССЗ у больного по данным от НСДАГ<sub>1–6</sub>, — и формирует диагноз.

В [5] были определены параметры НС:

- (1) алгоритм нечетких рассуждений Сугено;
- (2) пространство выходов  $Y = Y_1 \times \dots \times Y_{N_Y}$  — четких констант на интервале  $[0, 1]$ , размерность которого (число классов)  $N_Y = 7, 7, 10, 2, 6, 5, 19, 22, 2, 15$  соответственно;
- (3) пространство входов  $X = X_1 \times \dots \times X_{N_X}$  — нечеткие лингвистические переменные, размерность которого  $N_X = 30, 20, 80, 20, 65, 65, 190, 130, 25, 64$ ;
- (4) сигмоидная и треугольная функции принадлежности;
- (5) размеры баз знаний для НСДАГ<sub>1</sub>–НСДАГ<sub>9</sub>, НССС равны 10, 14, 15, 10, 30, 20, 30, 20, 24 и 70 соответственно и содержат нечеткие правила вида  $(x_{1j} = a_{1j}\Theta_j \dots \Theta_j x_{nj} = 1_{nj}) \Rightarrow y = d_j$ , где  $j$  — номер правила в базе знаний,  $\Theta_j$  — логическая операция, а  $d_j$  — четкое число (1 — истина, 0 — ложь);
- (6) дефазификация — модифицированный вариант метода центра тяжести для одноточечных множеств.

Функции принадлежности входных лингвистических переменных построены методом экспертного опроса [16], а пространства входов и выходов  $X, Y$  и база знаний КВ построены по рекомендациям комитета экспертов Всероссийского

**Таблица 3** Результаты исследования НС диагностики критериев оценки сердечно-сосудистого риска, вторичной АГ у пациента и оценки степени и стадии АГ, степени риска ССЗ

№	Наименование НС	Чувствительность НС	Специфичность НС	Средне-квадратическая ошибка классификации патологического состояния здоровья пациента
1	Диагностика поражений органов-мишеней	0,9	1	0,141
2	Диагностика факторов риска АГ	1	1	0
3	Диагностика цереброваскулярных болезней	0,85	0,8	0,1127
4	Диагностика метаболического синдрома и сахарного диабета	1	1	0
5	Диагностика заболеваний периферических артерий	0,95	0,95	0,187
6	Диагностика ишемической болезни сердца	0,95	0,9	0,061
7	Диагностика эндокринной АГ	0,8	0,85	0,135
8	Диагностика паренхиматозной нефропатии	0,75	0,75	0,164
9	Диагностика реноваскулярной АГ	1	1	0
10	Оценка степени и стадии АГ, степени риска ССЗ	1	1	0

научного общества кардиологов (ВНОК) [1], исследованиям Staessen с коллегами [17] и результатам анализа бесед с экспертом КОКБ.

Тестирование НС, решающих подзадачи «Диагностика критериев оценки сердечно-сосудистого риска и вторичной АГ у пациента» и подзадачу «Оценка степени и стадии АГ, степени риска ССЗ», проведено на 40 наблюдениях (табл. 3).

Для тестирования НС сформированы выборки из архива карт пациентов 1-го кардиологического отделения КОКБ размерностью  $N = 40$  наблюдений каждая. Таким образом, на каждую из девяти НС подавались по 20 наблюдений с диагнозом, содержащим диагностируемые патологии, и по 20 — не содержащие признаки болезни. По аналогии с нейросетями для каждой НС оценены чувствительность ( $Sen$ ) и специфичность ( $Sp$ ) и вычислены среднеквадратические ошибки классификации (2). Для каждой НС  $Sen \rightarrow 1$ ,  $Sp \rightarrow 1$  и  $(Sen - Sp) \rightarrow \rightarrow 0$ , а среднеквадратическая ошибка классификации патологического состояния

обследуемого либо равна нулю, либо не превышает 0,16, что повышает качество диагнозов Виртуального консилиума. Среднеквадратическая ошибка модели, оценивающей степень и стадию АГ, а также степень риска ССЗ, которая работает последней и фактически формирует диагноз, равна нулю, что значительно повышает достоверность диагноза системы. В итоге тестирование показало возможности включения всех моделей НСДАГ<sub>1</sub>–НСДАГ<sub>9</sub> и НССС в гетерогенное модельное поле.

### 3 Результаты экспериментов с лабораторным прототипом Виртуального консилиума для решения сложной задачи диагностики артериальной гипертензии

Наряду с гетерогенным модельным полем основная составляющая Виртуального консилиума — интерпретатор, алгоритм работы которого рассмотрен в [5]. Он имитирует коллективное принятие решений медицинским консилиумом, моделирует деятельность ЛПР (кардиолога) по организации консилиума, взаимодействию с врачами узких специализаций и формированию окончательного диагноза АГ у пациента.

Результат работы алгоритма — архитектура ВКДАГ — список Ls<sup>m</sup> как метод решения сложной задачи диагностики АГ (СЗДАГ).

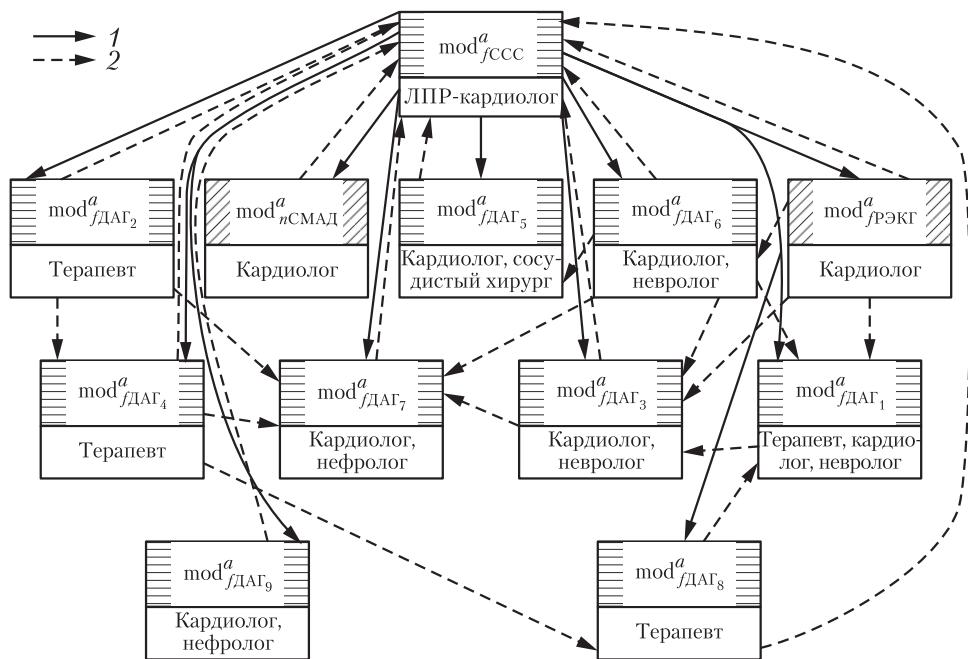
$$\begin{aligned} \text{Ls}^m = & \\ = & \left\{ R_{\text{int}}^{\text{upr}} \rightarrow \text{prb}_{\text{ППС}_{1-9}}^h \text{prb}_{\text{ДАГ}_{1-9}}^h \zeta_{\text{СЗДАГ}}^u \left( \text{mod}_{\text{ППС}_{1-9}}^a \rightarrow I^g, \text{mod}_{\text{ДАГ}_{1-9}}^a \rightarrow I^f \right), \right. \\ & R_{\text{int}}^{\text{dop}} \rightarrow \text{prb}_{\text{СМАД}}^h \text{prb}_{\text{CCC}}^h \zeta_{\text{СЗДАГ}}^u \left( \text{mod}_{\text{СМАД}}^a \rightarrow I^n, \text{mod}_{\text{CCC}}^a \rightarrow I^f \right), \\ & R_{\text{int}}^{\text{dop}} \rightarrow \text{prb}_{\text{ДАГ}_{1-9}}^h \text{prb}_{\text{CCC}}^h \zeta_{\text{СЗДАГ}}^u \left( \text{mod}_{\text{ДАГ}_{1-9}}^a \rightarrow I^f, \text{mod}_{\text{CCC}}^a \rightarrow I^f \right), \\ & \left. R_{\text{int}}^{\text{dop}} \rightarrow \text{prb}_{\text{РЭКГ}}^h \text{prb}_{\text{ДАГ}_1}^h \zeta_{\text{СЗДАГ}}^u \left( \text{mod}_{\text{РЭКГ}}^a \rightarrow I^n, \text{mod}_{\text{ДАГ}_1}^a \rightarrow I^f \right) \right\}. \end{aligned}$$

В Ls<sup>m</sup> знаки подзадач из декомпозиции СЗДАГ [4] (prb<sup>h</sup>) заменены на знаки выбранных из гетерогенного модельного поля моделей (mod<sup>a</sup>); prb<sup>h</sup><sub>РЭКГ</sub>, prb<sup>h</sup><sub>СМАД</sub>, prb<sup>h</sup><sub>ППС<sub>1-9</sub></sub>, prb<sup>h</sup><sub>ДАГ<sub>1-9</sub></sub>, prb<sup>h</sup><sub>CCC</sub> — задачи из декомпозиции неоднородной задачи СЗДАГ [4] соответственно: задача «Распознавание и интерпретация ЭКГ», задача «Анализ суточного мониторирования АД (СМАД)», группа из девяти технологических подзадач «Построение информативного набора признаков (симптомов) при диагностике заболеваний из областей 1–9», группа из 9 функциональных подзадач «Диагностика критериев оценки сердечно-сосудистого риска и вторичной АГ у пациента» и подзадача «Оценка степени и стадии АГ, степени риска ССЗ». Области 1–9 — это области однородных параметров, полученные редукцией сложной задачи, отображенные в декомпозиции СЗДАГ и содержащие подзадачи диагностики: поражений органов мишней (индекс 1), факторов

риска (2), цереброваскулярных болезней (3), метаболического синдрома и сахарного диабета (4), заболеваний периферических артерий (5), ишемической болезни сердца (6), эндокринной АГ (7), паренхиматозной нефропатии (8) и реноваскулярной АГ (9) соответственно. Каждой модели ( $mod^a$ ) сопоставлен интерпретатор  $I^v = \{I^g, I^f, I^n\}$  для имитации рассуждений специалиста, а каждому отношению интеграции сопоставлен интерфейс  $\zeta_{СЗДАГ}^u$  для обмена информацией.

На рис. 4 показано взаимодействие всех 12 функциональных моделей гетерогенного модельного поля, решающих задачи из декомпозиции СЗДАГ [4]; модели ЛПР (кардиолога) и моделей знаний врачей узких специализаций. Технологические элементы решают задачи безусловной оптимизации на множестве диагностируемых признаков и исключают из рассмотрения Виртуального консилиума избыточные и взаимозависимые симптомы заболеваний.

На рис. 4 показаны информационно-управляющие отношения дополнительности «эксперт–эксперт» и «эксперт–ЛПР» между моделями, по которым передаются промежуточные оценки вплоть до формирования ЛПР заключительного диагноза. Также между моделью ЛПР и моделями, содержащими знания экс-



**Рис. 4** Взаимодействие модели ЛПР (кардиолога) и моделей, содержащих знания экспертов в Виртуальном консилиуме: 1 — отношения сотрудничества «ЛПР–эксперт»; 2 — информационно-управляющие отношения дополнительности «эксперт–эксперт» и «эксперт–ЛПР»

пертов, показаны отношения сотрудничества «ЛПР–эксперт», формирующие организационную структуру ВКДАГ. Однаковой штриховкой на рис. 4 отмечены блоки — модели подзадач, решаемых посредством одной и той же методологии (НС — ДАГ<sub>1–9</sub> и ССС или нейросетей — РЭКГ и СМАД). В блоках указываются профессии врачей, знания и принятие решений которых модель имитирует при диагностике заболеваний областей 1–9 из декомпозиции СЗДАГ [4].

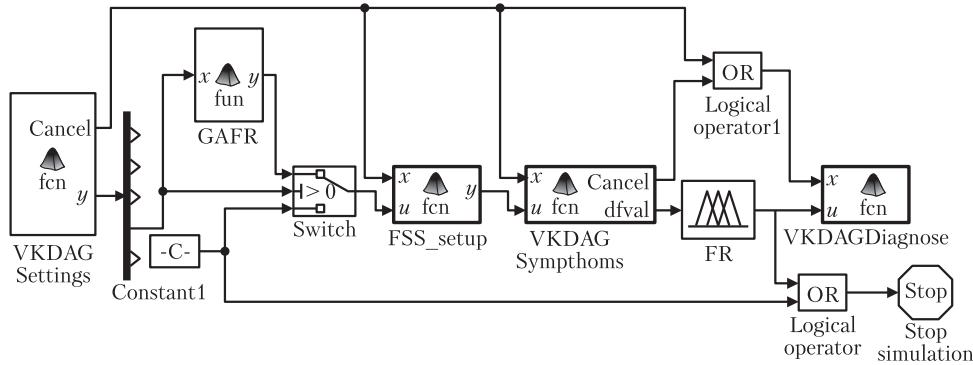
Архитектура Виртуального консилиума для диагностики АГ разработана и протестирована в системе визуального программирования Simulink — МАТ-LAB. Суть исследования алгоритма Виртуального консилиума состояла в следующем. В консилиуме принимали участие: кардиолог (в том числе и в качестве ЛПР), невролог, нефролог, терапевт, эндокринолог, уролог (см. рис. 4). Все эксперименты проводились на лабораторном экземпляре ВКДАГ, имеющем ограничения на возраст пациентов, на распознаваемые на ЭКГ патологии, на гипертоническую ретинопатию, на систолическую и некоторые вторичные АГ. Однако, так как гипертоническая болезнь или эссенциальная АГ, не связанная с вторичной АГ, встречается в 90%–95% случаев [1], то 20% вторичных АГ, не диагностируемых Виртуальным консилиумом, не влияют существенно на порог «замаскированных и недовыявленных» диагнозов у 70% пациентов амбулаторно-клинического учреждения в экспериментальных исследованиях.

Если в экспериментах с гетерогенным модельным полем тестировалась каждая его составляющая автономно и ее вход не обусловливался выходом других элементов поля, в экспериментах с ВКДАГ исследовалось проявление свойств дополнительности и сотрудничества, что имитировало коллективные рассуждения специалистов по постановке диагноза. Иными словами, цель экспериментов — проверить, с какой ошибкой модели, релевантные по отдельности, будут работать совместно через межмодульные интерфейсы ( $\zeta^u_{\text{СЗДАГ}}$ ), обмениваясь информацией между собой для заключительного диагноза.

Для лабораторных экспериментов в КОКБ по рекомендации эксперта (врача-кардиолога КОКБ) было отобрано 40 архивных амбулаторных карт. В них содержались как частные заключения врачей узких специализаций (невролог, нефролог, терапевт, эндокринолог, уролог) по состоянию здоровья объекта диагностики, так и результирующий диагноз — заключение ведущего врача-кардиолога. Исследовались заключения двух врачей каждой специализации, принадлежащих одной и той же медицинской школе, т. е. противоречий в извлеченных знаниях не было. Отсюда результаты экспериментов обусловлены знаниями врачей КОКБ, используемыми ими методиками диагностики, а также особенностями аппаратного обеспечения больницы.

### **3.1 Структура Виртуального консилиума для диагностики артериальной гипертензии в Simulink и межмодульные интерфейсы**

Архитектура Виртуального консилиума — модульная и гибкая для «сборки» гибридной интеллектуальной системы с учетом запроса пользователя. Создана



**Рис. 5** Модуль структуры Виртуального консилиума, интегрирующий интерфейс модификации структуры ВКДАГ, интерфейс ввода симптоматической информации, блок «Построение информативного набора признаков (симптомов) при диагностике заболеваний из области  $k$ », блок «Диагностика группы критериев оценки сердечно-сосудистого риска и вторичной АГ у пациента из области  $k$ » и блок «Оценка степени и стадии АГ, степени риска ССЗ»

библиотека прототипа ВКДАГ с блоками всех элементов гетерогенного модельного поля: ИНС, НС и ГА. Она состоит из девяти модулей (рис. 5), интегрирующих ГАППС $_k$ –НСДАГ $_k$ –НCCCC. На рис. 5 номер области из декомпозиции СЗДАГ [4]  $k = 2$  для примера НС диагностики факторов риска АГ (НСДАГ $_2$ ).

На рис. 5 блок VKDAGSettings открывает интерфейс пользователя (ЛПР, врача-кардиолога) для спецификации режима работы Виртуального консилиума: «Анализ СМАД» (решается посредством нейросетей, блок СМАД), «Распознавание ЭКГ» (решается посредством нейросетей, блок РЭКГ), «Построение информативного набора признаков (симптомов) при диагностике заболеваний из областей 1–9» (решается посредством ГА, блоки ППС $_1$ –ППС $_9$ ) и подзадачи из группы «Диагностика критериев оценки сердечно-сосудистого риска и вторичной АГ у пациента» (решаются посредством НС, блоки ДАГ $_1$ , ..., ДАГ $_9$ ).

Если ППС $_k$ ,  $k = \overline{1, 9}$ , выбрана для решения, то активируется блок «ППС $_k$ » (на рис. 5 это GAFR), с которого передается информация на блок «Модификация подсистем диагностики» (на рис. 5 это FSS\_setup) в соответствии с результатами решения технологических подзадач «Построение информативного набора признаков (симптомов) при диагностике заболеваний». Он корректирует работу блоков диагностики поражений органов мишеней, факторов риска, цереброваскулярных болезней, метаболического синдрома и сахарного диабета, заболеваний периферических артерий, ишемической болезни сердца, эндокринной АГ, паренхиматозной нефропатии и реноваскулярной АГ (блоки ДАГ $_1$ , ..., ДАГ $_9$  соответственно).

Далее активируется блок интерфейса пользователя (на рис. 5 это VKDAGSymptoms) и вводятся показатели состояния здоровья пациента (симптомы), передаваемые в блоки ДАГ<sub>1</sub>, …, ДАГ<sub>9</sub> (на рис. 5 это FR). Последние в свою очередь передают информацию о патологиях пациента блоку ССС (SSS), решающему подзадачу «Оценка степени и стадии АГ, степени риска ССЗ» (на рис. 5 это VKDAGDiagnose), и являются результатами моделирования принятия решения экспертами, врачами смежных специальностей — невропатологом, терапевтом, хирургом, эндокринологом, окулистом. При этом блок ССС моделирует принятие решения ЛПР — врачом-кардиологом.

В результирующей структуре ВКДАГ присутствует и блок NNECG, передающий информацию на блоки диагностики поражений органов-мишеней (РОМ, ДАГ<sub>1</sub>), цереброваскулярных болезней (CVB, ДАГ<sub>3</sub>) и ишемической болезни сердца (IBS, ДАГ<sub>6</sub>). Формат выхода блока NNECG совпадает с форматом входов блоков РОМ, CVB и IBS, поэтому не требует преобразований. Также в результирующей структуре ВКДАГ в Simulink есть блок NNSMAD, передающий информацию о требуемых нормальных значениях суточного АД по полу и возрасту пациентов на блок SSS.

### **3.2 Результаты тестирования алгоритма, имитирующего коллективные рассуждения над гетерогенным модельным полем**

На вход Виртуального консилиума для диагностики АГ подавалась тестовая выборка размерностью  $N = 40$  из архива карт пациентов 1-го кардиологического отделения КОКБ, состоящая из наблюдений  $(X_{11}, \dots, X_{1629}; X_{21}, \dots, X_{2128}; Y_1, \dots, Y_{96})$ , где  $X_{11}, \dots, X_{1629}$  — нечеткие лингвистические переменные;  $X_{21}, \dots, X_{2128}$  — вектор образа ЭКС из чисел интервала  $[0, 1]$  на основе записей кардиоциклов, соответствующих выделенным 6 патологиям и нормальному состоянию;  $Y_1, \dots, Y_{96}$  — четкие константы, принимающие значение 0 или 1, элементы диагноза: степень АГ; стадия эссенциальной АГ (или гипертонической болезни); класс общего риска ССЗ; факторы риска, поражения органов-мишеней и ассоциативные клинические состояния, присутствующие у пациента.

В экспериментах решение задачи диагностики АГ имитируется в соответствии с примером декомпозиции [4] (см. рис. 4). Сначала работают модели первой очереди — модели технологических элементов (ГАППС<sub>1–9</sub>), корректирующие множества входных переменных моделей НСДАГ<sub>1–9</sub> и НССС. Затем обработка информации передается функциональным элементам:

- (1) модели второй очереди «отправляют» информацию на модели третьей, пятой, шестой и седьмой очередей — ИНСРЭКГ (модель, решающая задачу распознавания ЭКГ), ИНССМАД (формирует оптимальные множества показателей суточного давления), НСДАГ<sub>9</sub>, НСДАГ<sub>2</sub> и НСДАГ<sub>6</sub>;
- (2) третья очередь содержит модели НСДАГ<sub>4</sub> и НСДАГ<sub>5</sub>, передающие выходную информацию на вход моделей четвертой и седьмой очередей;

- (3) четвертая очередь содержит модель НСДАГ<sub>8</sub>, передающую информацию модели пятой очереди (НСДАГ<sub>1</sub>), которая в свою очередь передает информацию НСДАГ<sub>3</sub> (шестая очередь);
- (4) от НСДАГ<sub>3</sub> передается информация НСДАГ<sub>7</sub> (седьмая очередь).

Последней работает модель НСДАГ<sub>7</sub>, формирующая заключительный диагноз, на вход которой передается выходная информация функциональных моделей второй–седьмой очередей.

По результатам работы Виртуального консилиума вычислялась среднеквадратическая ошибка постановки диагноза и время, затрачиваемое на обследование и формирование заключения о состоянии пациента:

- подсчитывалось отклонение диагноза от эталонного по каждому выходу ВКДАГ для  $N$  наблюдений в тестовой выборке, т. е.  $N$  ошибок классификации по  $l$ -му ( $l = \overline{1, N_k}$ ) выходу системы  $\delta_{lj}$  (равна 1, если совпадает с эталоном, и 0 — если нет), которые складывались и усреднялись;
- затем вычислялась среднеквадратическая ошибка постановки диагноза ( $f$ ) по формуле (2) при  $N = 40$  и  $N_k = 110$ .

По итогам лабораторных экспериментов в КОКБ прототип Виртуального консилиума диагностировал АГ со среднеквадратической ошибкой постановки диагноза  $f = 0,0837$ , т. е. давал верный диагноз в 80% случаев. В амбулаторно-клинических учреждениях диагноз в среднем не выявляется у 70% пациентов. При этом время, затрачиваемое врачом на постановку диагноза, складывается из времени процедуры обследования ( $t_o$ ), времени на обработку результатов ( $t_p$ ) и времени на запись и оформление заключения ( $t_z$ ) и составляет 15–20 мин. В экспериментах общее время диагностики АГ лабораторным прототипом уменьшилось примерно на 20%, а ( $t_z + t_p$ ) составило примерно 30 с.

По результатам лабораторных экспериментов с прототипом Виртуального консилиума в 1-м кардиологическом отделении КОКБ можно предположить, что ВКДАГ, доведенный до стадии практического применения в лечебно-профилактических учреждениях широкого профиля Калининградской области (фельдшерско-акушерских пунктах, амбулаториях, медико-санитарных частях и поликлиниках), позволит повысить качество постановки диагноза на 50% и сократить среднее время, затрачиваемое врачом на одного пациента, на 20%.

Также по результатам экспериментов с прототипом Виртуального консилиума среди студентов-медиков специальности «Лечебное дело» Балтийского федерального университета имени Иммануила Канта, проходящих интернатуру в КОКБ, можно сделать следующий вывод: Виртуальный консилиум — это хороший тренинг в постановке диагноза и анализе ЭКГ без участия преподавателя по курсу «Неотложная кардиология и функциональная диагностика». Работая с системой, студенты смогут повысить качество знаний, накапливая в памяти стереотипы принятия решений в области диагностики АГ. Скорость принятия ими качественного решения и степень уверенности во время приема пациентов в ходе

своей учебной и врачебной практики будут пропорциональны числу ситуаций, которые студенты отработают с применением Виртуального консилиума.

База знаний Виртуального консилиума обладает свойством универсальности, так как создана по всероссийским и европейским рекомендациям ВНОК для диагностики АГ [1, 18] и дополнена знаниями врачей КОКБ. При переносе системы в лечебные учреждения других регионов необходимо провести тестирование базы знаний системы на эмпирических данных (амбулаторных архивных картах) соответствующего учреждения, по результатам которого, возможно, потребуется модификация и/или дополнение базы знаний.

Таким образом, по результатам лабораторных экспериментов предполагается, что решение сложной задачи диагностики АГ посредством функциональной гибридной системы, разработанной в рамках проблемно-структурной (ПС) методологии, повысит качество клинической диагностики и снизит порог недовыявленных диагнозов симптоматической АГ и эссенциальной АГ с 70% до 20%, а также снизит время, затрачиваемое на обследование пациента и формирование диагноза.

#### 4 Заключение

Представлены результаты последних этапов разработки лабораторного прототипа Виртуального консилиума для диагностики АГ в рамках ПС-методологии как ФГиИС: имитации диагностики АГ на ЭВМ средствами пакета Simulink — MATLAB 7, модельные эксперименты и интерпретация их результатов. Разработаны алгоритмы для обучения и тестирования ИНСРЭКГ и ИНССМАД, рассчитаны оптимальные значения числа нейронов скрытых слоев сетей. Для обучения и тестирования ИНСРЭКГ также разработаны алгоритмы создания тестовых и обучающих выборок. Для ГАППС<sub>1–9</sub> (результат работы всех девяти ГА — это оптимальный набор входных признаков для НСДАГ<sub>1–9</sub>) создан и построен классический ГА ввиду особенностей функции приспособленности индивидуумов, вычисляемой с помощью разработанных НСДАГ<sub>1–9</sub>. Нечеткие системы исследованы и показали допустимые значения среднеквадратической ошибки классификации (0–0,187) и достаточную способность как обнаруживать заболевание (не ниже чем в 75% случаев), так и определять его отсутствие (не ниже чем в 75% случаев).

Применение методов гибридных интеллектуальных систем — искусственных гетерогенных систем — в медицинской информатике позволит снизить порог «замаскированных и недовыявленных» диагнозов симптоматической АГ и гипертонической болезни с 70% до 20%, что улучшит качество клинической диагностики АГ врачом-терапевтом. Открывается возможность снижения времени обследования и постановки диагноза, что актуально в условиях роста нагрузки на практикующих врачей.

Полученные результаты исследования лабораторного экземпляра ВКДАГ в КОКБ подтверждают эффективность предлагаемого подхода для проекти-

рования медицинских диагностических систем как гетерогенных искусственных диагностических систем со свойствами дополнительности, сотрудничества и относительности знаний, синтезирующих интегрированные методы и модели, разнообразие которых устраниет разнообразие диагностической информации об организме человека.

## **Литература**

1. Плавунов Н. Ф., Степанова В. В., Караполова Ю. Л., Котовская Ю. В., Кобалова Ж. Д., Мусеев В. С. Роль артериальной гипертонии в структуре и исходах госпитализаций многопрофильного скоропомощного стационара // Артериальная гипертензия, 2004. Т. 10. Вып. 3. С. 30–36.
2. ESH/ESC. Guidelines for the management of arterial hypertension: The Task Force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC) // J. Hypertens., 2013. Vol. 31. No. 7. P. 1281–1357. doi: 10.1097/01.hjh.0000431740.32696.cc.
3. Колесников А. В., Кириков И. А. Методология и технология решения сложных задач методами функциональных гибридных интеллектуальных систем. — М.: ИПИ РАН, 2007. 387 с.
4. Кириков И. А., Колесников А. В., Румовская С. Б. Исследование сложной задачи диагностики артериальной гипертензии в методологии искусственных гетерогенных систем // Системы и средства информатики, 2013. Т. 23. № 2. С. 96–114. doi: 10.14357/08696527130208.
5. Кириков И. А., Колесников А. В., Румовская С. Б. Функциональная гибридная интеллектуальная система для поддержки принятия решений при диагностике артериальной гипертензии // Системы и средства информатики, 2014. Т. 24. № 1. С. 153–179. doi: 10.14357/08696527140110.
6. Зубдинов Ю. И. Азбука ЭКГ. — 3-е изд. — Ростов-на-Дону: Феникс, 2003. 169 с.
7. Петрухин В. А., Мамаев В. Н., Дроздов Р. В. Использование сравнительного анализа последовательных векторэлектрокардиограмм для оценки локальных изменений миокарда // Комп'ютерні засоби, мережі та системи, 2005. № 4. С. 131–137.
8. Аль Мабрук М. Аппаратно-программные средства и алгоритмы распознавания патологий сердца на основе персепtronных сетей. Дисс. . . канд. техн. наук. — Владимир, 2011. 203 с.
9. The research source for complex physiologic signals // PhysioNet. <http://physionet.org>.
10. Исаков Р. В., Кузнецов А. А., Сушкина Л. Т. Методы и средства обработки информации при оценке функционального состояния организма человека. — Владимир: Собор, 2006. 144 с.
11. Рагайян Р. М. Анализ биомедицинских сигналов / Пер. с англ. под ред. А. П. Немирко. — М.: Физматлит, 2007. 440 с. (Rangayyan R. M. Biomedical signal analysis: A case-study approach. — Wiley, 2002. 516 p.)
12. Хайкин С. Нейронные сети: полный курс / Пер. с англ. — 2-е изд., испр. — М.: Вильямс, 2006. 1110 с. (Haykin S. Neural networks: A comprehensive foundation. — 2nd ed. — Singapore: Pearson Prentice Hall, 1999. 823 p.)
13. Колесников А. В. Гибридные интеллектуальные системы. Теория и технология разработки / Под ред. А. М. Яшина. — СПб.: СПбГТУ, 2001. 711 с.

14. Колесников А. В., Кириков И. А., Румовская С. Б. и др. Разработка системы с базой знаний для поддержки принятия решения при диагностике артериальной гипертензии: Отчет о НИР. — Калининград: Калининградский филиал ИПИ РАН, 2010. 135 с.
15. Новикова В. А., Андреева Е. Ю., Туйкина Д. К. Искусственный интеллект и экспертные системы. — Курск: КГУ, 2004. 45 с.
16. Штровба С. Д. Проектирование нечетких систем средствами Matlab. — М.: Горячая линия – Телеком, 2007. 288 с.
17. Staessen J., O'Brien E., Atkins N., Bulpitt C., Cox J., Fagard R., O'Malley K., Thijs L., Amery A. The increase in blood pressure with age and body mass index is overestimated by conventional sphygmomanometry // Am. J. Epidemiol., 1992. Vol. 136. P. 450–459.
18. Комитет экспертов ВНОК. Диагностика и лечение артериальной гипертензии. Российские рекомендации // Системные гипертензии, 2010. Вып. 3. С. 5–26.

Поступила в редакцию 27.04.14

---

## EXPERIMENTAL RESEARCH OF LABORATORY PROTOTYPE OF AN ARTIFICIAL HETEROGENEOUS SYSTEM FOR DIAGNOSIS OF ARTERIAL HYPERTENSION

I. A. Kirikov<sup>1</sup>, A. V. Kolesnikov<sup>1,2</sup>, and S. B. Rumovskaya<sup>1</sup>

<sup>1</sup>Kaliningrad Branch of the Institute of Informatics Problems, Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation

<sup>2</sup>Immanuel Kant Baltic Federal University, 14 Nevskogo Str., Kaliningrad 236041, Russian Federation

**Abstract:** The paper contains the results of imitation of Virtual medical consultation for arterial hypertension diagnostics (functional hybrid system) by means of MATLAB 7 — Simulink and model experiments. The paper introduces the results of testing models from heterogeneous model field: functional modules for electrocardiogram recognition, for investigation of the module analyzing the data of daily monitoring of arterial pressure, and for classification of a patient's state of health; and also, technological modules for experimental utility estimation of symptomatic information for making a diagnosis by means of genetic algorithms. Experiments with laboratory prototype showed the opportunity of decreasing the threshold of "disguised and negotiating diagnosis" of arterial hypertension from 70% to 20% by using it in practice in medical and preventive treatment facilities of wide profile.

**Keywords:** laboratory prototype of virtual medical consultation; arterial hypertension; neural networks; genetic algorithms; fuzzy systems; functional hybrid intelligent system

**DOI:** 10.14357/08696527140309

## References

1. Plavunov, N. F., V. V. Stepanova, Yu. L. Karaulova, Yu. V. Kotovskaya, Zh. D. Kobalava, and V. S. Moiseev. 2004. Rol' arterial'noy gipertonii v strukture i ishodakh gospitalizatsiy mnogoprofil'nogo skoropomoshchnogo statsionara [The role of arterial hypertension in the structure and outcomes of hospitalizations in multidisciplinary first-aid hospital]. *Arterial'naya Gipertenziya* [Arterial Hypertension] 10(3):30–36.
2. ESH/ESC. 2013. Guidelines for the management of arterial hypertension: The Task Force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *J. Hypertens.* 31(7):1281–1357. doi: 10.1097/HJH.0000431740.32696.cc.
3. Kolesnikov, A. V., and I. A. Kirikov. 2007. *Metodologiya i tekhnologiya resheniya slozhnykh zadach metodami funktsional'nykh gibriddenykh intellektual'nykh sistem* [Methodology and technology for solving of complex problems using the methodology of functional hybrid artificial systems]. Moscow: IPI RAN. 387 p.
4. Kirikov, I. A., A. V. Kolesnikov, and S. B. Rumovskaya. 2013. Issledovanie slozhnoy zadachi diagnostiki arterial'noy gipertenzii v metodologii iskusstvennykh heterogenykh sistem [Research of the complex problem at diagnosing of the arterial hypertension within the methodology of artificial heterogeneous systems]. *Sistemy i Sredstva Informatiki—Systems and Means of Informatics* 23(2):96–114. doi: 10.14357/08696527130208.
5. Kirikov, I. A., A. V. Kolesnikov, and S. B. Rumovskaya. 2014. Funktsional'naya gibridnaya intellektual'naya sistema dlya podderzhki prinyatiya resheniya pri diagnostike arterial'noy gipertenzii [Functional hybrid intelligent decision support system for diagnosing of the arterial hypertension]. *Sistemy i Sredstva Informatiki—Systems and Means of Informatics* 24(1):153–179. doi: 10.14357/08696527140110.
6. Zudbinov, Yu. I. 2003. *Azbuka ECG* [Alphabet of ECG]. 3rd ed. Rostov-na-Donu: Phoenix Publ. 169 p.
7. Petrukhin, V. A., V. N. Mamaev, and R. V. Drozdov. 2005. Ispol'zovanie sravnitel'nogo analiza posledovatel'nykh vektorelektrokardiogramm dlya otsenki lokal'nykh izmeneniy miokarda [The comparative analysis of sequential vectorelectroretinography for assessment of local changes of the myocardium]. *Komp'yuterni zasoby, merezhi ta systemy* [Computer means, networks, and systems] 4:131–137.
8. Al' Mabruk, M. 2011. Apparatno-programmnye sredstva i algoritmy raspoznavaniya patologiy serdtsa na osnove perseptronnykh setey [Hardware, software and algorithms for recognition of heart pathology based on perceptron-type networks]. D.Sc. Thesis. Vladimir. 203 p.
9. The research source for complex physiologic signals. Available at: <http://physionet.org> (accessed February 27, 2014).
10. Isakov, R. V., A. A. Kuznetsov, and L. T. Sushkova. 2006. *Metody i sredstva obrabotki informatsii pri otsenke funktsional'nogo sostoyaniya organizma cheloveka* [Methods and means of information processing in the assessment of the functional state of human body]. Vladimir: Sobor. 144 p.
11. Rangayyan, R. M. 2002. *Biomedical signal analysis: A case-study approach*. Wiley. 516 p.
12. Haykin, S. 1999. *Neural networks: Comprehensive foundation*. 2nd ed. Singapore: Pearson Prentice Hall. Inc. 823 p.

13. Kolesnikov, A. V. 2001. Gibridnye intellektual'nye sistemy. *Teoriya i tehnologiya razrabotki* [Hybrid artificial systems. Theory and development technology]. St. Petersburg: SPbGTU. 710 p.
14. Kolesnikov, A. V., I. A. Kirikov, S. B. Rumovskaja, et al. 2010. Razrabotka sistemy s bazoy znaniy dlya podderzhki prinyatiya resheniya pri diagnostike arterial'noy gipertenzii [Development of knowledge-based system for decision support at diagnosing of arterial hypertension]. Report. Kaliningrad: Kaliningrad Branch of IPI RAN. 135 p. [In Russian.]
15. Novikova, V. A., E. Yu. Andreeva, and D. K. Tuykina. 2004. *Iskusstvennyy intellekt i ekspertnye sistemy* [Artificial intelligence and expert systems]. Kursk: KGU. 45 p.
16. Shtovba, S. D. 2007. *Proektirovanie nechetkikh sistem sredstvami Matlab* [Designing of fuzzy systems with Matlab tools]. Moscow: Goryachaya liniya – Telekom. 288 p.
17. Staessen, J., E. O'Brien, N. Atkins, C. Bulpitt, J. Cox, R. Fagard, K. O'Malley, L. Thijs, and A. Amery. 1992. The increase in blood pressure with age and body mass index is overestimated by conventional sphygmomanometry. *Am. J. Epidemiol.* 136(4):450–459.
18. Komitet ekspertov VNOK. 2010. Diagnostika i lechenie arterial'noy gipertenzii. Rossiyskie rekomendatsii [Diagnosing and treatment of arterial hypertension. Russian recommendations]. *Sistemnye Gipertenzii* [Systemic Hypertension] 3:5–26.

Received April 27, 2014

## Contributors

**Kirikov Igor A.** (b. 1955) — Candidate of Science (PhD) in technology, Director, Kaliningrad Branch of the Institute of Informatics Problems, Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation; baltbipiran@mail.ru<sup>1</sup>

**Kolesnikov Alexander V.** (b. 1948) — Doctor of Science in technology, Director, professor, Immanuel Kant Baltic Federal University, 14 Nevskogo Str., Kaliningrad 236041, Russian Federation; senior scientist, Kaliningrad Branch of the Institute of Informatics Problems, Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation; avkolesnikov@yandex.ru

**Rumovskaya Sophiya B.** (b. 1985) — programmer I, Kaliningrad Branch of the Institute of Informatics Problems, Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation; sophiyabr@gmail.com

## ОСОБЕННОСТИ ФОРМИРОВАНИЯ НОМЕНКЛАТУРЫ И КОЛИЧЕСТВА КОМПОНЕНТОВ ЗИП В АВТОМАТИЗИРОВАННЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ В ЗАЩИЩЕННОМ ИСПОЛНЕНИИ

*А. А. Зацаринный<sup>1</sup>, С. В. Козлов<sup>2</sup>, А. И. Гаранин<sup>3</sup>*

**Аннотация:** Рассмотрены организационно-методические проблемы обеспечения надежности комплексов средств автоматизации (КСА) автоматизированных информационных систем в защищенном исполнении (АИС ЗИ) на основе анализа, обобщения и оценки опыта работ по техническому надзору за их эксплуатацией. Показаны факторы, определяющие время подготовки и проведения ремонта средств электронно-вычислительной техники. Обоснованы предложения по реализации иерархического принципа формирования номенклатуры и количества компонентов в составе ЗИП (запасные части, инструменты, принадлежности) в виде четырехуровневой системы комплектования ЗИП и их использованию в ходе организации работ по ремонту технических средств (ТС). Представлены расчеты уровня запаса запасных частей в составе ЗИП на объекте. Дан пример оценки количества ТС в составе ЗИП для обеспечения требуемого показателя достаточности (ПД) ЗИП.

**Keywords:** автоматизированная система в защищенном исполнении; надежность; отказ технического средства; иерархическая система и номенклатура ЗИП; восстановление работоспособности; показатель достаточности ЗИП

**DOI:** 10.14357/08696527140310

### 1 Введение

Анализ, обобщение и оценка опыта работ по техническому надзору за АИС ЗИ показывают, что выбор подходов к обеспечению их надежности и, прежде всего, ремонтопригодности ТС во многом ограничивается требованиями по защите информации.

Так, в случае отказа ТС из состава КСА АИС ЗИ его ремонт при наличии в составе ЗИП аналога отказавшего элемента производится на объекте эксплуатации либо доставляется в ремонтный орган. В первом варианте время восстановления отказавшего ТС определяется временем выполнения следующих работ:

---

<sup>1</sup>Институт проблем информатики Российской академии наук, azatsarinny@ipiran.ru

<sup>2</sup>Институт проблем информатики Российской академии наук, sv\_kozlov@mail.ru

<sup>3</sup>Институт проблем информатики Российской академии наук, algaranin@mail.ru

- диагностика ТС с выявлением отказавшего оборудования;
- доставка оборудования из мест хранения ЗИП к месту эксплуатации ТС и замена отказавшего элемента на работоспособный из состава ЗИП;
- проверка работоспособности отремонтированного ТС;
- проведение специальных исследований (СИ) отремонтированного ТС специализированной организацией и отправка в экспертную организацию (ЭО) результатов СИ и проекта нового предписания на его эксплуатацию;
- проведение экспертизы в ЭО и доведение ее результатов.

Во втором варианте время ремонта отказавших изделий АИС ЗИ дополнительно включает временные затраты на следующие работы:

- доставка неисправных ТС в ремонтную организацию, их дефектация;
- приобретение компонентов для ремонта, проведение их специальной проверки (СП), СИ отремонтированного ТС и экспертиза их результатов;
- передача отремонтированного компонента на объект применения.

В соответствии с основными положениями теории надежности [1, 2] время, не связанное непосредственно с отысканием причины отказа, его устранением и проверкой работоспособности отремонтированного оборудования, обычно относят к так называемым «организационным простоям». В АИС ЗИ, в отличие от АИС в обычном исполнении, к организационным простоям относится время, затрачиваемое на проведение СП закупленных для замены неисправных компонентов и проведение СИ отремонтированного оборудования специализированной организацией и получение заключения (предписания на эксплуатацию) по их результатам от ЭО. Чаще всего время организационных простоеv с учетом времени доставки неисправных ТС в ремонтную организацию и обратно может достигать 2–3 мес. Поэтому поиск путей сокращения времени восстановления работоспособности функциональных изделий за счет совершенствования организации подготовки и проведения ремонта является актуальной задачей.

## **2 Предложения по сокращению времени восстановления функциональных изделий на основе уточнения подходов к формированию номенклатуры ЗИП**

В состав одиночных комплектов ЗИП объектов АИС обычно включают следующие комплектующие: системные платы, накопители на жестких магнитных дисках (НЖМД), модули памяти, сетевые карты, источники бесперебойного питания (ИБП) и т. п.

Для сокращения времени восстановления работоспособности отказавших функциональных изделий представляется целесообразным изменить номенклатуру комплектующих в составе ЗИП [3]. Исходя из опыта проведенных работ по поддержанию работоспособности АИС ЗИ, предлагается в составе ЗИП предусматривать следующие группы оборудования:

**Группа 1** — отдельные функциональные изделия (серверные ЭВМ, комплексы ЭВМ для рабочих станций, телекоммуникационные средства и средства защиты информации), прошедшие СП и СИ и имеющие соответствующие регистрационные документы. При этом перед включением в состав ЗИП серверов и рабочих станций они должны быть загружены общим и специальным программным обеспечением (ПО).

При использовании оборудования этой группы для восстановления работоспособности КСА АИС ЗИ время восстановления ТС включает:

- время, необходимое на организацию получения и доставку оборудования из мест хранения ЗИП в места эксплуатации отказавшего ТС;
- время включения оборудования в состав КСА АИС ЗИ;
- время проверки работоспособности ТС в составе КСА.

Замена на объекте отказавшего функционального изделия на исправное осуществляется службой эксплуатации с временем восстановления до 2 ч.

После проведения ремонта отказавшее ТС возвращается обратно в состав функционального изделия. Отремонтированное функциональное изделие по решению службы эксплуатации либо используется для восполнения ЗИП, либо возвращается на прежнее место функционирования, а заменившее его функциональное изделие из состава ЗИП возвращается обратно в состав ЗИП. В обоих указанных случаях время проведения ремонта отказавшего функционального изделия (с учетом времени, необходимого на проведение СИ и получение заключения от ЭО) можно считать временем пополнения состава ЗИП. Опыт работ по техническому надзору за АИС ЗИ показывает, что ЗИП обычно пополняется 1 раз в год.

Таким образом, использование оборудования ЗИП первой группы позволяет сократить время восстановления функциональных изделий АИС ЗИ с 2–3 мес. до 3–4 ч, а период пополнения ЗИП — с 1–2 лет до 2–3 мес.

**Группа 2** — отдельные компоненты функциональных изделий (системные блоки ЭВМ, мониторы, периферийные устройства (принтеры и сканеры) из состава автоматизированных рабочих мест (АРМ), клавиатура, ИБП и др.), прошедшие СП.

Вторая группа оборудования ЗИП используется для замены отказавших ТС в составе функциональных изделий, замененных из ЗИП оборудованием первой группы. При использовании оборудования второй группы для восстановления работоспособности ТС АИС ЗИ время восстановления функционального изделия включает:

- время замены отказавшего оборудования на исправное из состава ЗИП, включая его доставку из мест хранения, время включения оборудования в схему изделия и его загрузку ПО, и может достигать 2–4 ч;

- время на организацию и проведение СИ отремонтированного изделия, которое может составлять от 2–3 дней до 1–2 нед.;
- время получения заключения по результатам СИ от ЭО, которое может составлять до 2–3 мес.

Таким образом, использование оборудования ЗИП второй группы позволяет обеспечить готовность оборудования ЗИП первой группы после его использования для ремонта функционального изделия АИС ЗИ и возвращения в ЗИП неисправного оборудования, что позволяет исключить время организационных простоев, связанных с проведением СИ отремонтированного функционального изделия. Время восстановления изделий АИС ЗИ при этом может сократиться с 2–3 мес. до 3–4 ч, а период пополнения ЗИП — с 1–2 лет до 2–3 мес.

**Группа 3** — отдельные узлы, блоки и платы компонентов функциональных изделий, в том числе:

- материнские платы, видеокарты, прошедшие СП и тематические исследования встроенного ПО микросхем BIOS (Basic Input/Output System);
- НЖМД, платы памяти, контроллеры, встроенные блоки питания и другие устройства, прошедшие СП.

Третья группа оборудования ЗИП предназначена для замещения компонентов в составе отказавших ТС, замененных из ЗИП оборудованием второй группы.

При использовании оборудования этой группы для восстановления работоспособности ТС время восстановления функционального изделия включает те же затраты времени, что и при использовании оборудования второй группы и может составить до 2–3 мес.

**Группа 4** — расходные материалы (аккумуляторные батареи, картриджи и др.), прошедшие СП.

Таким образом, наиболее весомый вклад в сокращение времени восстановления работоспособности функциональных изделий в АИС ЗИ обеспечивается за счет использования для ремонта оборудования ЗИП первой группы. Вместе с тем в составе ЗИП необходимо иметь все четыре группы оборудования и компонентов. Так, необходимость включения в состав ЗИП оборудования второй и третьей групп объясняется следующими основными причинами.

1. После восстановления работоспособности функционального изделия, отказавшие ТС из его состава могут быть восстановлены службой эксплуатации объекта с использованием оборудования ЗИП второй и третьей групп. При их отсутствии в составе ЗИП неисправное оборудование необходимо отправлять в ремонтные органы, а это требует дополнительного времени.
2. Срок службы АИС ЗИ обычно составляет не менее 10 лет. В то же время в связи с быстрым развитием средств вычислительной техники номенклатура

выпускаемых промышленностью комплектующих для ремонта ТС (системные платы, процессоры, память и др.) обновляется каждые 2–3 года [4]. Это означает, что через 2–3 года после поставки оборудования на объекты АИС ЗИ в ремонтных органах может не оказаться нужной номенклатуры комплектующих и будет необходимо заменять их современными аналогами. В этом случае для системных плат и других элементов, содержащих устройства BIOS, потребуется проводить их тематические исследования, что также связано с дополнительными временными и финансовыми затратами.

Оценку количественных характеристик комплектования ЗИП с учетом предложенных групп оборудования ЗИП целесообразно производить в соответствии с методикой расчета запасов в комплекте ЗИП, изложенной в [4, 5].

### **3 Основные положения методики расчета уровня запаса запасных частей в комплекте ЗИП**

Рассматривается задача проектирования комплектов ЗИП АИС ЗИ, обеспечивающих заданный уровень ПД при минимальных затратах.

Показатель достаточности [6] системы ЗИП характеризует снижение надежности пары «объект – конкретная система ЗИП» по сравнению с надежностью пары «объект – бесконечная система ЗИП» («бесконечная система ЗИП» предполагает, что необходимые запасные элементы не закончатся в системе ЗИП до окончания срока эксплуатации объекта). Показатель достаточности, по существу, — коэффициент готовности комплекта ЗИП  $K_{\text{г.зип}}$  (вероятность того, что система ЗИП исправна) [5].

Оценим необходимое количество запасных элементов в ЗИП-О объекта при различных вариантах его комплектования, с тем чтобы при минимуме затрат на ЗИП выполнялось неравенство:

$$K_{\text{г.зип}} \geq K_{0.\text{зип}} = \frac{R_0}{K_\infty},$$

где  $K_{0.\text{зип}}$  — требуемый коэффициент готовности системы ЗИП;  $R_0$  — требуемое значение коэффициента готовности объекта с учетом реальной системы ЗИП;  $K_\infty$  — коэффициент готовности объекта при бесконечной системе ЗИП.

Отказом системы ЗИП условно называется такое состояние пары «объект – система ЗИП», при котором объект полностью или частично потерял работоспособность из-за отказа одного из составляющих его элементов, а система ЗИП не может предоставить нужного запасного элемента. Расчет запасов в комплекте ЗИП состоит из расчетов запасов каждого типа в отдельности и последующей оценки ПД и суммарных затрат по комплекту ЗИП в целом. Необходимые запасы элементов в комплекте ЗИП зависят от принятой в системе стратегии их пополнения, среди которых рассмотрим [5]:

- периодическое пополнение — запас элементов данного типа восстанавливается через заранее заданные интервалы времени  $T_{\text{пп}} = T_{\text{pi}}$ ;
- непрерывное пополнение — стратегию непрерывного пополнения применяют для запасов восстанавливаемых элементов, которые либо пополняются из ЗИП более высокого уровня, либо возвращаются после ремонта в ЗИП, из которого были изъяты. Эта стратегия пополнения характеризуется средним временем ремонта одного элемента данного типа  $T_{\text{пп}} = T_{\text{pi}}$ .

Приемлемая точность вычисления ПД комплектов ЗИП в целом по характеристикам отдельных запасов обеспечивается при условии, что требуемые значения ПД комплекта удовлетворяют неравенству:

$$K_{\text{г.зип}} = \prod_{i=1}^N K_{\text{зги}} \geq 0,9,$$

где  $K_{\text{зги}}$  — коэффициент готовности запасных частей  $i$ -го типа в комплекте ЗИП;  $i$  — порядковый номер типа запаса в комплекте ЗИП ( $i = 1, \dots, N$ );  $N$  — общее число типов запасных частей.

Исходные данные для расчета запасов в комплекте ЗИП АИС ЗИ:

$m_i$  — количество составных частей  $i$ -го типа в изделии;

$\lambda_{zi}$  — интенсивность замен составных частей  $i$ -го типа в изделии ( $\lambda_{zi} = 1/T_{oi}$ , где  $T_{oi}$  — средняя наработка на отказ  $i$ -го элемента);

$n_{io}$  — необходимый уровень запаса элементов  $i$ -го типа в комплекте ЗИП;

$\varepsilon_0$  — необходимая точность вычисления ПД комплекта ЗИП.

Значение ПД комплекта ЗИП вычисляется по формуле [7]

$$K_{\text{г.зип}} = \exp \left\{ \prod_{i=1}^N R_{io}(n_{io}, a_{io}) \right\}, \quad (1)$$

где  $a_{io}$  — среднее число поступающих в комплект ЗИП заявок на запасные части для запаса каждого типа за период пополнения:

$$a_{io} = m_i \lambda_{zi} T_{\text{пп}}; \quad (2)$$

$R_{io}(n_{io}, a_{io})$  — промежуточный расчетный показатель, который при периодическом пополнении ЗИП с периодом пополнения  $T_{ni}$  вычисляется по формуле [7]:

$$R_{io}(n_{io}, a_{io}) = -\ln \left\{ 1 - \frac{1}{a_{io}} \left[ e^{-a_{io}} \sum_{\gamma=n_i+2}^{\infty} (\gamma - n_{io} - 1) \frac{a_{io}^\gamma}{\gamma!} \right] \right\}. \quad (3)$$

**Примечание.** При расчетах по формуле (3) выражение в квадратных скобках, содержащее бесконечную сумму, вычисляют до такого значения индекса суммирования  $\gamma$ , при котором это выражение впервые удовлетворяет неравенству

$$[a_{io}, n_{io}, \gamma] \leq \frac{\varepsilon_0}{2N}. \quad (4)$$

При непрерывном пополнении запасов ЗИП-О с периодом пополнения  $T_{pi}$  промежуточный расчетный показатель вычисляется по формуле [7]:

$$R_{io}(n_{io}, a_{io}) = -\ln \left[ 1 - \frac{a_{io}^{n_{io}+1}}{(n_{io}+1)! \sum_{k=0}^{n_{io}+1} a_{io}^k / k!} \right]. \quad (5)$$

#### 4 Пример расчета количественных характеристик вариантов комплектования ЗИП

Пусть объект состоит из  $n$  серверов и  $m$  рабочих станций. Кроме количественного состава известна  $T_{oi}$  — средняя наработка на отказ оборудования  $i$ -го типа в каждой группе оборудования. Предложенный состав объекта позволит рассмотреть наиболее типовые варианты комплектования ЗИП.

Исходные данные по составу оборудования объекта, учитываемые при оценке состава ЗИП, представлены в табл. 1.

По первому варианту — принятая система комплектования ЗИП, период пополнения запасов элементов данного типа  $T_{nn} = 8760$  ч (1 год). В состав ЗИП включаются системные платы, НЖМД, модули памяти, сетевые карты, ИБП, клавиатура, конвертеры и т. п.

По второму варианту принята 4-уровневая система комплектования.

Исходные данные для расчетов первого варианта представлены в табл. 2 (колонки 1–4). Используя выражение (2), определим  $a_{io}$  и заполним колонку 6 табл. 2. Значение промежуточного расчетного показателя  $R_{io}(n_{io}, a_{io})$  вычислим по формуле (3). Численное значение выражения в квадратных скобках в (3) при фиксированном  $i$  зависит только от  $n_{io}$ . Для получения заданной точности вычисления ПД принимаем  $\varepsilon_0 = 0,1$  (в соответствии с (4)  $\varepsilon_0/2N = 0,005$ ) и путем итераций, начиная с  $n_{io} = 1$  и увеличивая  $n_{io}$  на каждом шаге на единицу, отыскиваем такое значение  $n_{io}$ , при котором выполняется неравенство (4). Полученные таким образом значения  $n_{io}$  и  $R_{io}(n_{io}, a_{io})$  вставляем в табл. 2 (колонки 5 и 7 соответственно).

Анализ полученных результатов (колонка 5) показывает, что для обеспечения заданных показателей надежности функционирования  $n = 6$  серверов с ПД не

**Таблица 1** Исходные данные по составу оборудования объекта

Оборудование	$i$	Обозначение	$T_{0i}$ , ч	$m_i$ , шт.	Наименование оборудования
Первой группы	1	1-1	30 000	6	Сервер
	2	1-2	12 000	15	АРМ
Второй группы	1	2-1	15 000	15	Системный блок
	2	2-2	25 000	15	Монитор
	3	2-3	18 000	15	Клавиатура
	4	2-4	18 000	15	Манипулятор мышь
	5	2-5	12 000	15	ИБП
Третьей группы	1	3-1	14 000	6	Сменный модуль питания сервера 750 Вт
	2	3-2	18 000	6	Модуль памяти сервера
	3	3-3	18 000	6	Системная плата сервера
	4	3-4	20 000	6	Сетевая карта сервера
	5	3-5	10 000	6	НЖМД сервера
	6	3-6	20 000	6	Средство контроля доступа
	7	3-7	30 000	6	Процессор сервера
	8	3-8	18 000	15	Системная плата ЭВМ
	9	3-9	18 000	15	Процессор ЭВМ
	10	3-10	18 000	15	Модуль памяти ЭВМ
	11	3-11	10 000	15	НЖМД ЭВМ
	12	3-12	20 000	15	Сетевая карта ЭВМ
	13	3-13	20 000	15	Видеокарта ЭВМ
	14	3-14	12 000	15	Сменный модуль питания ЭВМ 450 Вт

менее 0,9 и при периодическом пополнении состава ЗИП 1 раз в год потребуется иметь в составе ЗИП объекта от 6 до 12 запасных элементов различного типа. Для обеспечения заданных показателей надежности функционирования  $m = 15$  АРМ потребуется иметь в составе ЗИП объекта от 12 до 23 запасных элементов различного типа.

Результирующее значение показателя  $R_{io}(n_{io}, a_{io})$  представлено в нижней строке табл. 2. Подставляя это значение в выражение (1), получим значение ПД комплекта ЗИП-О  $K_{\text{г.зип}} = 0,957$ . Другими словами, при указанном в колонке 5 количестве запасных частей в составе ЗИП коэффициент готовности системы ЗИП составит  $K_{\text{г.зип}} = 0,957$ .

Во втором варианте для пополнения оборудования ЗИП 1-й группы и части оборудования 2-й группы (оборудование, которое может быть отремонтировано на объекте путем замены отказавших комплектующих из ЗИП третьей группы, например системный блок ЭВМ) используется отремонтированное оборудование, и поэтому принятая стратегия непрерывного пополнения с числовым параметром пополнения  $T_{np} = 1440$  ч (2 мес.), а для оборудования 3-й группы и оставшегося оборудования 2-й группы — периодическая стратегия пополнения с параметром  $T_{i,1} = 8760$  ч.

**Таблица 2** Исходные данные и результаты расчетов при первом варианте комплектования ЗИП ( $T_{\text{пп}} = 8760$  ч)

$i$	Обозначение	$m_i$ , шт.	$\lambda_{zi}$ , 1/ч	$n_{io}$ , шт.	$a_{io}$	$R_{io}(n_{io}, a_{io})$
1	2	3	4	5	6	7
1	3-1	6	0,000071	10	3,73176	0,00098899
2	3-2	6	0,000056	7	2,94336	0,00579369
3	3-3	6	0,000056	7	2,94336	0,00579369
4	3-4	6	0,00005	7	2,628	0,00320029
5	3-5	6	0,0001	12	5,256	0,00170599
6	3-6	6	0,00005	7	2,628	0,00320029
7	3-7	6	0,000033	6	1,73448	0,0012477
8	3-8	15	0,000056	15	7,3584	0,00203253
9	3-9	15	0,000056	15	7,3584	0,00203253
10	3-10	15	0,000056	15	7,3584	0,00203253
11	3-11	15	0,0001	23	13,14	0,00213117
12	3-12	15	0,00005	14	6,57	0,00175696
13	3-13	15	0,00005	14	6,57	0,00175696
14	3-14	15	0,000083	20	10,9062	0,00208838
15	2-2	15	0,00004	12	5,256	0,00170599
16	2-3	15	0,000056	15	7,3584	0,00203253
17	2-4	15	0,000056	15	7,3584	0,00203253
18	2-5	15	0,000083	20	10,9062	0,00208838
$\sum_{io=1}^{N_o} R_{io}(n_{io}, a_{io})$						<b>0,04362113</b>

Исходные данные и результаты расчета комплекта ЗИП при втором варианте его формирования представлены в табл. 3 (колонки 1–5).

Результаты вычислений с использованием выражения (5) представлены в колонке 6 — для оборудования первой группы (сервер и АРМ) в строках 1–2, для оборудования второй группы (системный блок ЭВМ) — строка 1. Для оборудования 3-й группы результаты расчетов такие же, как и в табл. 2.

Результирующее значение промежуточного показателя  $R_{io}(n_{io}, a_{io})$  представлено в нижней строке табл. 3. Подставляя это значение в выражение (3), получим значение ПД комплекта ЗИП  $K_{\text{г.зип-о}} = 0,9499$ .

Из табл. 3 следует, что для обеспечения заданных показателей надежности функционирования 6 серверов и 15 АРМ потребуется иметь в составе ЗИП объекта 2 сервера, 6 АРМ и 5 системных блоков ЭВМ. При этом на порядок сокращается время восстановления неисправного оборудования.

При расчетах минимизация затрат на создание комплекта ЗИП обеспечивалась включением в состав ЗИП минимального количества запасов каждого типа, при которых обеспечивается заданный уровень ПД.

**Таблица 3** Исходные данные и результаты расчетов при втором варианте комплектования ЗИП

Оборудование	$i$	Обозначение	$m_i$ , шт.	$\lambda_{zi}$ , 1/ч	$T_i$ , ч	$n_i$ , шт.	$a_{io}$	$R_{io}(n_{io}, a_{io})$
0	1	2	3	4	5	6	7	8
Первой группы	1	1-1	6	0,000033	1440	2	0,28512	0,0029096
	2	1-2	15	0,000083	1440	6	1,7928	0,0019695
Второй группы	1	2-1	15	0,000067	1440	5	1,4472	0,00300821
	2	2-2	15	0,000040	8760	12	5,256	0,00170599
	3	2-3	15	0,000056	8760	15	7,3584	0,00203253
	4	2-4	15	0,000056	8760	15	7,3584	0,00203253
	5	2-5	15	0,000083	8760	20	10,9062	0,00208838
Третьей группы	1	3-1	6	0,000071	8760	10	3,73176	0,00098899
	2	3-2	6	0,000056	8760	7	2,94336	0,00579369
	3	3-3	6	0,000056	8760	7	2,94336	0,00579369
	4	3-4	6	0,00005	8760	7	2,628	0,00320029
	5	3-5	6	0,0001	8760	12	5,256	0,00170599
	6	3-6	6	0,00005	8760	7	2,628	0,00320029
	7	3-7	6	0,000033	8760	6	1,73448	0,0012477
	8	3-8	15	0,000056	8760	15	7,3584	0,00203253
	9	3-9	15	0,000056	8760	15	7,3584	0,00203253
	10	3-10	15	0,000056	8760	15	7,3584	0,00203253
	11	3-11	15	0,0001	8760	23	13,14	0,00213117
	12	3-12	15	0,00005	8760	14	6,57	0,00175696
	13	3-13	15	0,00005	8760	14	6,57	0,00175696
	14	3-14	15	0,000083	8760	20	10,9062	0,00208838
$\sum_{io=1}^{N_o} R_{io}(n_{io}, a_{io})$							<b>0,05150844</b>	

Таким образом, предложенная иерархическая система комплектования ЗИП с учетом особенностей АИС ЗИ за счет рационального подхода к формированию номенклатуры компонентов ЗИП позволяет сократить в 8–12 раз время, обусловленное организационными простоями в процессе организации и проведения ремонта ТС в составе АИС ЗИ. Этот подход к формированию номенклатуры ЗИП апробирован в ходе выполнения работ по обеспечению функционирования и поддержания работоспособности ведомственной АИС ЗИ. В дальнейшем в развитие иерархической системы комплектования ЗИП АИС ЗИ целесообразно провести ее технико-экономическую оценку.

## Литература

- Гнеденко Б. В., Беляев Ю. К., Соловьев А. Д. Математические методы в теории надежности. — М.: Наука, 1965. 524 с.

2. Каиштанов В. А., Медведев А. И. Теория надежности сложных систем. — М.: Физматлит, 2010. 606 с.
3. Синицын И. Н., Шаламов А. С. Лекции по теории систем интегрированной логистической поддержки. — М.: ТОРУС ПРЕСС, 2012. 624 с.
4. Зацаринный А. А., Гаранин А. И., Козлов С. В., Кондрашев В. А. Особенности расчета комплектов ЗИП в автоматизированных информационных системах в защищенном исполнении // Системы и средства информатики, 2013. Т. 23. № 1. С. 113–132.
5. Беляев Ю. К., Богатырев В. А., Болотин В. В. и др. Надежность технических систем: Справочник / Под ред. проф. И. А. Ушакова. — М.: Радио и связь, 1985. 608 с.
6. ГОСТ РВ 0015.705-2008 СРПП ВТ. Запасные части, инструменты и принадлежности. Основные положения. — М.: Изд-во стандартов, 2008. 22 с.
7. ГОСТ РВ 27.3.03-2005. Надежность военной техники. Оценка и расчет запасов в комплектах ЗИП. — М.: Стандартинформ, 2005. 40 с.

*Поступила в редакцию 22.08.14*

---

## **THE PECULIARITIES OF NOMENCLATURE AND SPARE COMPONENTS QUANTITY COMPOSITION FOR SECURED AUTOMATED INFORMATION SYSTEMS**

*A. A. Zatsarinniy, S. V. Kozlov, and A. I. Garanin*

Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str.,  
Moscow 119333, Russian Federation

**Abstract:** Organizational and methodological problems of providing reliability for automation means complexes of secured automated information systems based on analysis, generalization, and evaluation of experience of supervisions on their operations are considered. The article shows some factors which determine lead-up and operational time of maintenance and repairs concerning electronic components. The suggestions regarding the structure of the spare parts list and the components quantity and their utilizations during repair work organization as like as four-level hierarchical structure are justified by the authors. Calculations of stock necessity of spare parts at the object are presented. Finally, there is an example of real calculations regarding quantity of necessary spare parts stock to achieve conditions of reliability.

**Keywords:** Secured Automated Information System; reliability; hardware failures; hierarchical principles; spare parts list; system recovery; sufficiency of spare parts parameter

**DOI:** 10.14357/08696527140310

## References

1. Gnedenko, B. V., Y. K. Belyaev, and A. D Soloviev. 1965. *Matematicheskie metody v teorii nadezhnosti* [Mathematics methods in reliability theory]. Moscow: Nauka. 524 p.
2. Kashtanov, V. A., and A. I. Medvedev. 2010. *Teoriya nadezhnosti slozhnykh sistem* [The reliability theory for complex systems] Moscow: Fizmatlit. 606 p.
3. Sinitsyn, I. N., and A. S. Shalamov. 2012. *Lektsii po teorii sistem integrirovannoy logisticheskoy podderzhki* [Lectures on theory of integrated logistic support systems]. Moscow: TORUS PRESS. 624 p.
4. Zatsarinny, A. A., A. I. Garanin, S. V. Kozlov, and V. A. Kondrashev. 2013. Osnobennosti rascheta komplektov ZIP v avtomatizirovannykh informatsionnyh sistemakh v zashchishchennom ispolnenii [About some particularities of spare parts sets calculation for secured information systems]. *Sistemy i Sredstva Informatiki—Systems and Means of Informatics* 23(1):113–132.
5. Belyaev, Y. K., V. A. Bogatyrev, V. V. Bolotin, et al. 1985. *Nadezhnost' tekhnicheskikh system: Spravochnik* [Reliability of technical systems: Reference book]. Ed. A. I. Ushakov. Moscow: Radio i Svyaz'. 608 p.
6. GOST RV 0015.705-2008 SRPP VT. 2001. Zapasnye chasti, instrumenty i prinadlezhnosti. Osnovnye polozheniya. [Spare parts, tools, and accessories. Main theses]. Moscow: Izdatel'stvo Standartov Publ. 22 p.
7. GOST RV 27.3.03-2005. 2007. Nadezhnost' voennoy tekhniki. Otsenka i raschet zapasov v komplektakh ZIP [Reliability of military technics. The evaluation and calculation of stock necessity by spare sets]. Moscow: StandartInform Publ. 40 p.

Received August 22, 2014

## Contributors

**Zatsarinnyy Alexander A.** (b. 1951) — Doctor of Science in technology, professor, Deputy Director, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; azatsarinny@ipiran.ru

**Kozlov Sergey V.** (b. 1955) — Candidate of Science (PhD) in technology, Head of Department, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sv\_kozlov@mail.ru

**Garanin Alexander I.** (b. 1951) — Candidate of Science (PhD) in technology, senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; algaranin@mail.ru

## ОСОБЕННОСТИ РАЗРАБОТКИ ДОКУМЕНТАЦИИ ДЛЯ АВТОМАТИЗИРОВАННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ В ЗАЩИЩЕННОМ ИСПОЛНЕНИИ

*А. А. Зацаринный<sup>1</sup>, Ю. С. Ионенков<sup>2</sup>*

**Аннотация:** Рассмотрены особенности разработки документации для автоматизированных информационных систем (АИС) в защищенном исполнении с учетом требований существующей нормативно-технической базы, а также опыта внедрения ряда систем на объектах заказчика. Определены основные системотехнические и организационные проблемы, возникающие при определении состава документации для АИС. Предложен системный подход к представлению состава документации для АИС, являющегося достаточным для проектирования, развертывания АИС, а также обеспечения ее эксплуатации с требуемым качеством. Определены особенности для АИС в защищенном исполнении. Предложен конкретный перечень рабочей документации, разрабатываемый при создании автоматизированных систем. Показано, что предлагаемые подходы достаточно эффективны при освоении документации обслуживающим персоналом. Излагаемые в статье результаты успешно апробированы на практике создания и внедрения нескольких АИС.

**Ключевые слова:** автоматизированная информационная система; рабочая, конструкторская, эксплуатационная документация

**DOI:** 10.14357/08696527140311

### 1 Введение

Современные АИС представляют собой взаимоувязанную (согласованную) совокупность трех составляющих [1–3]:

- (1) системотехнических решений (требования, состав, физическая и логическая структура, профиль протоколов и пр.);
- (2) технологий и специального программного обеспечения, реализующих функциональные задачи АИС;
- (3) комплексов технических средств (КТС), которые обеспечивают создание объектовых комплексов системы.

---

<sup>1</sup>Институт проблем информатики Российской академии наук, azatsarinny@ipiran.ru

<sup>2</sup>Институт проблем информатики Российской академии наук, uionenkov@ipiran.ru

Все три указанные составляющие в ходе создания АИС описываются в проектной и рабочей (рабочей конструкторской) документации. От номенклатуры и качества документации зависят сроки и качество выполнения работ по созданию системы, а также возможности освоения и дальнейшей эксплуатации системы пользователями и обслуживающим персоналом. Документация должна полностью отражать технические решения, реализованные в созданной системе и ее комплексах, обладать достаточной простотой для независимого от изготовителя освоения системы обслуживающим персоналом. Особую значимость приобретает качество документации при создании систем в защищенном исполнении. В последнее время в связи с сокращением времени на разработку систем и комплексов качество разрабатываемой документации снижается.

Целью данной статьи является рассмотрение проблем, подходов и особенностей разработки документации для АИС в защищенном исполнении на основе имеющегося у авторов опыта разработки подобных систем.

## 2 Основные проблемы разработки документации

В настоящее время процесс разработки АИС осуществляется на основе следующих документов:

- ГОСТ РВ класса 15.ххх — «Система разработки и постановки продукции на производство. Военная техника»;
- ГОСТ класса 34 — «Информационная технология. Комплекс стандартов на автоматизированные системы».

При создании АИС по ГОСТу РВ класса 15.ххх рабочая конструкторская документация разрабатывается на основе стандартов единой системы конструкторской документации (ЕСКД) — ГОСТ класса 2.ххх и единой системы программной документации (ЕСПД) — ГОСТ класса 19.ххх.

Номенклатура конструкторских документов при этом определяются ГОСТом 2.102-68 — «Виды и комплектность конструкторских документов» [4].

Виды, комплектность и правила выполнения эксплуатационных документов определяются ГОСТом 2.601-95 — «Эксплуатационные документы» [5].

Конструкторские и эксплуатационные документы оформляются в соответствии с ГОСТом 2.105-95 — «Общие требования к текстовым документам» [6].

Номенклатура программных документов определяется ГОСТом 19.101-77 — «Виды программ и программных документов» [7]. Требования к содержанию и оформлению конкретных программных документов определяются рядом ГОСТов класса 19.

При создании АИС по ГОСТ класса 34 рабочая документация наряду с использованием стандартов ЕСКД и ЕСПД разрабатывается с применением ряда документов из серии стандартов класса 34, таких как:

- ГОСТ 34.201-89 — «Виды, комплектность и обозначение документов при создании автоматизированных систем» [8];

- РД 50-34.698-90 — «Автоматизированные системы. Требования к содержанию документов» [9].

ГОСТ 34.201-89 определяет номенклатуру и обозначение документации, разрабатываемой на различных стадиях работы, при этом наименование и обозначение ряда документов отличаются от принятых в ЕСКД. РД 50-34.698-90 определяет содержание всех документов. Программная документация разрабатывается в соответствии с требованиями ЕСПД.

При создании автоматизированных систем в защищенном исполнении наряду с ГОСТом класса 34 применяется ГОСТ Р 51583-2000 — «Защита информации. Порядок создания автоматизированных систем в защищенном исполнении. Общие положения» [10], в котором раскрывается типовое содержание работ по защите информации применительно ко всем стадиям и этапам разработки АИС, предусмотренным ГОСТом класса 34.

Таким образом, имеющиеся нормативные документы достаточно полно определяют номенклатуру, содержание и порядок разработки документации.

Вместе с тем опыт проектирования АИС показывает, что в плане документирования их разработки имеются следующие проблемы:

- практическое отсутствие проектной и рабочей документации (имеется лишь документация, входящая в состав покупных аппаратных и программных средств);
- недостаточность документации как по номенклатуре, так и по полноте описания, что затрудняет освоение системы и ее эксплуатацию обслуживающим персоналом;
- несбалансированность документации, включая недостаточность описания отдельных подсистем и комплексов, в том числе подсистемы защиты информации.

Кроме того, полная номенклатура документации, приведенная в указанных выше ГОСТах, разрабатывается крайне редко, так как ряд документов на практике оказывается невостребованным.

Таким образом, при проектировании систем основными проблемными вопросами являются:

- определение совместно с заказчиком требований к системе в соответствии с нормативными документами;
- определение перечня документации в соответствии со спецификой системы;
- апробация разработанной документации на этапах предварительных испытаний и опытной эксплуатации.

### **3 Основные подходы к разработке документации**

В ходе создания АИС разрабатывается проектная и рабочая (рабочая конструкторская) документация. Проектная документация разрабатывается на эта-

пах эскизного и технического проекта и описывает проектные решения по системе и ее составным частям, включая решения по КТС, информационному, лингвистическому программному обеспечению, организационной структуре и т. п. Рабочая (рабочая конструкторская) документация разрабатывается на этапе «Рабочая документация» и включает в себя конструкторскую и эксплуатационную документацию, позволяющую создать и эксплуатировать АИС.

Многолетний опыт разработки АИС различного назначения в защищенном исполнении показывает, что наиболее целесообразно выделять следующие группы рабочей (рабочей конструкторской) документации:

- общесистемная документация;
- документация на комплексы и отдельные аппаратно-программные средства;
- документация на комплекс средств защиты информации;
- программная документация;
- проектно-сметная документация.

Такая структуризация всего комплекта документации позволяет обеспечить системное представление разрабатываемого изделия с его декомпозицией на составные части, а также наиболее удобное использование документации пользователями и обслуживающим персоналом.

Необходимо отметить, что все указанные выше группы документации разрабатываются применительно к конкретным комплексам. Для крупных же АИС, включающих ряд комплексов, разрабатываются лишь общесистемная документация, включая решения по подсистеме защиты информации (модель нарушителя, модель угроз, политика безопасности) и программная документация, которые содержат общие решения по построению системы и описывают ее основные функции.

Общесистемная документация содержит назначение системы (комплекса), ее функции, описание ее элементов и взаимосвязей с другими системами (комплексами). В состав общесистемной документации входят общее описание системы, схема деления, программа и методика испытаний системы.

Общее описание системы содержит назначение системы, ее описание, описание взаимосвязей с другими автоматизированными системами и при необходимости — описание подсистем.

Схема деления определяет состав изделия, входимость составных частей, их назначение и взаимосвязь. Уровень деления (раскрытия) изделия на составные части зависит от сложности и специфики изделия и устанавливается разработчиком по согласованию с заказчиком. Схема деления в соответствии с ГОСТом имеет статус обязательного конструкторского документа.

Программа и методика испытаний представляют собой организационно-методический документ, устанавливающий объект, цели, место и сроки испытаний, виды, последовательность и объем проводимых проверок, а также подробное описание практических действий, используемых при проведении испытаний.

При разработке АИС по ГОСТу РВ класса 15.xxx программа и методика испытаний оформляются в соответствии с ГОСТом РВ 15.211-2002 [11], а при использовании ГОСТа класса 34 оформление документа производится в соответствии с РД 50-34.698-90.

Документация на комплексы и отдельные аппаратно-программные средства включает в себя конструкторские и эксплуатационные документы.

К конструкторским документам относят графические и текстовые документы, определяющие состав и устройство изделия и содержащие необходимые данные для его разработки, контроля и приемки. В большинстве случаев разрабатываются следующие документы: спецификация (спецификация оборудования); схемы (структурные, функциональные, электрические и др.); таблицы (соединений, подключений); чертеж общего вида.

Эксплуатационные документы предназначены для эксплуатации изделий, ознакомления с их конструкцией, изучения правил эксплуатации, отражения сведений, удостоверяющих гарантированные изготовителем значения основных параметров и свойств изделия, гарантii и сведений по его эксплуатации. Представляется целесообразным разрабатывать следующий комплект эксплуатационных документов: руководство по эксплуатации (руководство пользователя, инструкция по эксплуатации КТС); формуляр (паспорт), ведомость комплекта запасных частей, инструмента и принадлежностей (ЗИП); ведомость потребности в материалах; ведомость эксплуатационных документов.

Основное содержание конструкторских и эксплуатационных документов приведено в табл. 1.

Указанная номенклатура документов используется при разработке в составе АИС объектовых комплексов, комплексов серверных и телекоммуникационных, комплексов средств отображения информации, автоматизированных рабочих мест различного назначения. Она вполне обеспечивает потребности разработчиков изделий и позволяет освоить и в последующем эксплуатировать изделия обслуживающему персоналу.

При создании АИС в защищенном исполнении в составе изделия разрабатывается комплекс средств защиты информации. Для указанного комплекса в целом разрабатывается такой же комплект конструкторских и эксплуатационных документов, как и для других комплексов. К особенностям комплекта документов на комплекс средств защиты информации можно отнести разработку таких документов, как руководство администратора безопасности и инструкция по антивирусной защите.

В ходе создания АИС в большинстве случаев разрабатываются комплексы специального программного обеспечения (СПО), реализующие функциональные задачи соответствующей системы. Для каждого из комплексов СПО разрабатывается программная документация в соответствии с ГОСТом класса 19. Номенклатура программных документов определяется ГОСТом 19.101-77 [7]. В состав программной документации целесообразно включать: спецификацию; текст программы; описание программы; формуляр; ведомость эксплуатационных

**Таблица 1** Содержание конструкторских и эксплуатационных документов

Наименование документа	Содержание документа	
	ЕСКД	ГОСТ 34.201-89
Конструкторские документы		
Спецификация	Состав документации, оборудования и программных средств изделия	—
Спецификация оборудования	—	Наименование и характеристики элементов изделия (ГОСТ 21.110-95)
Схема электрическая структурная	Составные части изделия и их взаимосвязи (ГОСТ 2.701-2008)	—
Схема структурная КТС	—	Состав КТС и связи между техническими средствами
Таблица соединений	Электрические соединения между монтажными изделиями	—
Таблица соединений и подключений	—	Электрические соединения между монтажными изделиями и подключение проводок к ним
Чертеж общего вида	<ul style="list-style-type: none"> <li>– Компоновка и расположение оборудования;</li> <li>– виды на плоскости электрических и иных проводок;</li> <li>– схема расположения шкафов или панелей;</li> <li>– перечень приборов, монтажных изделий и материалов</li> </ul>	<ul style="list-style-type: none"> <li>– Компоновка и расположение оборудования;</li> <li>– виды на плоскости электрических и иных проводок;</li> <li>– схема расположения шкафов или панелей;</li> <li>– перечень приборов, монтажных изделий и материалов</li> </ul>
Эксплуатационные документы		
Руководство по эксплуатации	<ul style="list-style-type: none"> <li>– Описание и работа;</li> <li>– использование по назначению;</li> <li>– техническое обслуживание;</li> <li>– текущий ремонт;</li> <li>– хранение;</li> <li>– транспортирование;</li> <li>– утилизация</li> </ul>	—
Руководство пользователя	—	<ul style="list-style-type: none"> <li>– Введение;</li> <li>– назначение и условия применения;</li> <li>– подготовка к работе;</li> <li>– описание ситуаций;</li> <li>– аварийные ситуации;</li> <li>– рекомендации по освоению</li> </ul>

*Продолжение табл. 1 на с. 162*

**Таблица 1** (*продолжение*) Содержание конструкторских и эксплуатационных документов

Наименование документа	Содержание документа	
	ЕСКД	ГОСТ 34.201-89
Инструкция по эксплуатации КТС	—	<ul style="list-style-type: none"><li>– Общие положения;</li><li>– структура КТС;</li><li>– средства вычислительной техники;</li><li>– аппаратура передачи данных</li></ul>
Формуляр	<ul style="list-style-type: none"><li>– Общие указания;</li><li>– основные сведения об изделии;</li><li>– основные технические данные;</li><li>– индивидуальные особенности изделия;</li><li>– комплектность;</li><li>– ресурсы, сроки службы и хранения;</li><li>– консервация;</li><li>– свидетельство об упаковывании;</li><li>– свидетельство о приемке;</li><li>– движение изделия при эксплуатации;</li><li>– учет работы изделия;</li><li>– учет технического обслуживания;</li><li>– работы при эксплуатации;</li><li>– хранение;</li><li>– ремонт;</li><li>– особые отметки</li></ul>	<ul style="list-style-type: none"><li>– Общие сведения;</li><li>– основные характеристики;</li><li>– комплектность;</li><li>– свидетельство о приемке;</li><li>– гарантии производителя;</li><li>– сведения о рекламациях</li></ul>
Паспорт	<ul style="list-style-type: none"><li>– Основные сведения о документе и технические данные;</li><li>– комплектность;</li><li>– ресурсы, сроки службы и хранения;</li><li>– консервация;</li><li>– свидетельство об упаковывании;</li><li>– свидетельство о приемке;</li><li>– движение изделия при эксплуатации;</li><li>– особые отметки</li></ul>	<ul style="list-style-type: none"><li>– Общие сведения;</li><li>– основные характеристики;</li><li>– комплектность;</li><li>– свидетельство о приемке;</li><li>– гарантитное обслуживание;</li><li>– сведения о состоянии;</li><li>– сведения о рекламациях</li></ul>

*Окончание табл. 1 на с. 163*

**Таблица 1** (окончание) Содержание конструкторских и эксплуатационных документов

Наименование документа	Содержание документа	
	ЕСКД	ГОСТ 34.201-89
Ведомость ЗИП	– Запасные части; – инструмент; – принадлежности; – материалы	—
Ведомость потребности в материалах	—	Наименование и характеристики запасных материалов и принадлежностей для изделия (ГОСТ 21.110-95)
Ведомость эксплуатационных документов	Перечень эксплуатационных документов согласно ГОСТу 2.601- 95	Перечень эксплуатационных документов согласно ГОСТу 34.201-89

документов; описание применения; руководство системного программиста; руководство оператора; программу и методику испытаний. Основное содержание программных документов приведено в табл. 2.

При создании АИС в защищенном исполнении должны проводиться тематические исследования комплексов СПО на отсутствие недекларированных возможностей, в том числе программных закладок. В интересах специализированной организации, проводящей тематические исследования, могут разрабатываться и другие программные документы, в частности пояснительная записка и руководство по сборке программного обеспечения.

Проектно-сметная документация разрабатывается применительно к конкретным объектам и включает:

- рабочий проект;
- сметную документацию.

Рабочий проект на объект, как правило, включает пояснительную записку, схемы соединений, схемы подключений, схемы расположения оборудования, схемы кабельных трасс и т. п. При выполнении рабочего проекта используются следующие нормативные документы:

- выбор параметров защитного заземления осуществляется в соответствии с правилами устройства электроустановок (ПУЭ);
- при подготовке помещений и определении плотности размещения аппаратно-программных средств в рабочих помещениях применяются соответствующие положения строительных и санитарных норм и правил;
- правила строительства кабельных трасс, а также кабельных вводов в здания регламентируются руководствами, стандартами и другими документами Министерства связи и массовых коммуникаций РФ;

**Таблица 2** Содержание программных документов

Наименование документа	Содержание документа
Спецификация (ГОСТ 19.202-78)	<ul style="list-style-type: none"><li>– Документация;</li><li>– комплексы;</li><li>– компоненты</li></ul>
Текст программы (ГОСТ 19.401-78)	Запись программы с необходимыми комментариями
Описание программы (ГОСТ 19.402-78)	<ul style="list-style-type: none"><li>– Общие сведения;</li><li>– функциональное назначение;</li><li>– описание логической структуры;</li><li>– используемые технические средства;</li><li>– вызов и загрузка;</li><li>– входные данные;</li><li>– выходные данные</li></ul>
Формуляр (ГОСТ 19.501-78)	<ul style="list-style-type: none"><li>– Общие указания;</li><li>– общие сведения;</li><li>– основные характеристики;</li><li>– комплектность;</li><li>– периодический контроль основных характеристик при эксплуатации и хранении;</li><li>– свидетельство о приемке;</li><li>– свидетельство об упаковке и маркировке;</li><li>– гарантийные обязательства;</li><li>– сведения о рекламациях;</li><li>– сведения о хранении;</li><li>– сведения о закреплении программного изделия при эксплуатации;</li><li>– сведения об изменениях;</li><li>– особые отметки</li></ul>
Ведомость эксплуатационных документов (ГОСТ 19.507-78)	Перечень программных эксплуатационных документов
Описание применения (ГОСТ 19.502-78)	<ul style="list-style-type: none"><li>– Назначение программы;</li><li>– условия применения;</li><li>– описание задачи;</li><li>– входные и выходные данные</li></ul>
Руководство системного программиста (ГОСТ 19.503-79)	<ul style="list-style-type: none"><li>– Общие сведения о программе;</li><li>– структура программы;</li><li>– настройка программы;</li><li>– проверка программы;</li><li>– дополнительные возможности;</li><li>– сообщения системному программисту</li></ul>

*Окончание табл. 2 на с. 165*

**Таблица 2** (окончание) Содержание программных документов

Наименование документа	Содержание документа
Руководство оператора (ГОСТ 19.505-79)	<ul style="list-style-type: none"> <li>– Назначение и условия применения программы;</li> <li>– характеристика программы;</li> <li>– обращение к программе;</li> <li>– входные и выходные данные;</li> <li>– сообщения</li> </ul>
Программа и методика испытаний (ГОСТ 19.301-79)	<ul style="list-style-type: none"> <li>– Объект испытаний;</li> <li>– цель испытаний;</li> <li>– состав предъявляемой документации;</li> <li>– технические требования;</li> <li>– порядок проведения испытаний;</li> <li>– методы испытаний</li> </ul>

– планы, схемы, чертежи и прочая графическая часть документации выполняется в соответствии со стандартами серии «Система проектной документации для строительства» (СПДС) — ГОСТ Р класса 21.xxx. Наиболее часто для этой цели применяется ГОСТ Р 21.1703-2000 — «Правила выполнения рабочей документации проводных средств связи» [12].

Затраты на монтажные и пусконаладочные работы основного оборудования, а также проектирование и оборудование инженерных систем (электроснабжения, вентиляции и кондиционирования, охранно-пожарной сигнализации, пожаротушения), обеспечивающих функционирование объектовых комплексов АИС, определяются на основе разработки сметной документации.

Сметная документация разрабатывается в соответствии с «Методикой определения стоимости строительной продукции на территории Российской Федерации МДС81-35-2004», введенной в действие Постановлением Госстроя от 5 марта 2004 г. № 15/1. Методика имеет в своем составе общие методические положения по составлению сметной документации и определению сметной стоимости строительства, выполнения ремонтных, монтажных и пусконаладочных работ на всех стадиях разработки проектной документации, формированию договорных цен и проведению расчетов за выполненные работы.

В некоторых случаях с целью унификации аппаратно-программных средств АИС разрабатываются опытные образцы типовых комплексов и автоматизированных рабочих мест различного назначения. При их разработке и испытаниях применяется ГОСТ РВ класса 15.xxx. Соответственно, документация на указанные комплексы разрабатывается на основе стандартов ЕСКД.

## **4 Заключение**

Состав и содержание рабочей документации, разрабатываемой при создании АИС, в значительной степени определяется спецификой создаваемой системы. В то же время в большинстве случаев прослеживается определенная общность подходов.

В статье представлен возможный подход к разработке документации для АИС в защищенном исполнении, отмечены основные проблемы и особенности разработки документации. Предложен конкретный перечень документации, разрабатываемый при создании автоматизированных систем. Данный перечень является минимально необходимым для качественной разработки АИС. Вместе с тем имеющийся у авторов опыт разработки АИС показывает, что данный комплект документации при надлежащем его исполнении является вполне достаточным для создания системы, ее освоения обслуживающим персоналом и эффективной эксплуатации.

## **Литература**

1. Зацаринный А. А., Ионенков Ю. С. Некоторые аспекты выбора технологии построения информационно-телекоммуникационных сетей // Системы и средства информатики, 2007. Вып. 17. С. 5–16.
2. Зацаринный А. А., Ионенков Ю. С. Методика выбора технических средств для построения телекоммуникационных сетей // Системы и средства информатики, 2009. Доп. вып. С. 4–14.
3. Зацаринный А. А., Ионенков Ю. С., Козлов С. В. Некоторые вопросы проектирования информационно-телекоммуникационных систем. — М.: ИПИ РАН, 2010. 218 с.
4. ГОСТ 2.102-68. Единая система конструкторской документации. Виды и комплектность конструкторских документов. — М.: Стандартинформ, 2002.
5. ГОСТ 2.601-95. Единая система конструкторской документации. Эксплуатационные документы. — М.: Стандартинформ, 2004.
6. ГОСТ 2.105-95. Единая система конструкторской документации. Общие требования к текстовым документам. — М.: Стандартинформ, 2002.
7. ГОСТ 19.101-77. Единая система программной документации. Виды программ и программных документов. — М.: Стандартинформ, 2010.
8. ГОСТ 34.201-89. Комплекс стандартов на автоматизированные системы. Виды, комплектность и обозначение документов при создании автоматизированных систем. — М.: Стандартинформ, 2005.
9. РД 50-34.698-90. Комплекс стандартов на автоматизированные системы. Автоматизированные системы. Требования к содержанию документов. — М.: Стандартинформ, 2005.
10. ГОСТ Р 51583-2000. Защита информации. Порядок создания автоматизированных систем в защищенном исполнении. Общие положения. — М.: Госстандарт России, 2003.

11. ГОСТ РВ 15.211-2002. Система разработки и постановки продукции на производство. Военная техника. Порядок разработки программ и методик испытаний опытных образцов изделий. — М.: Госстандарт России, 2004.
12. ГОСТ Р 21.1703-2000. Система проектной документации для строительства. Правила выполнения рабочей документации проводных средств связи. — М.: Госстрой России, 2004.

*Поступила в редакцию 23.07.14*

---

---

## THE PECULIAR PROPERTIES OF DOCUMENTATION DEVELOPMENT FOR SECURED AUTOMATED INFORMATION SYSTEMS

*A. A. Zatsarinny and Y. S. Ionenkov*

Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str.,  
Moscow 119333, Russian Federation

**Abstract:** The peculiar properties of documentation development for secured automated information systems (AIS) considering requirements of existing regulatory and technical base and experience of systems installation at customer's sites are described. The main system-technical and organizational problems concerning content of secured AIS documentation arising during implementation are determined. The paper suggests a systematic approach to presentation of secured AIS documentation content that is sufficient for development, deployment, and operation of secured AIS with necessary quality. The paper determines specialties of secured AIS and offers exact list of papers of working documentation developed for AIS. Finally, it is shown that suggested approaches are sufficiently effective for staff education. The results are tested in practice in the process of creation and installation of several secured AIS.

**Keywords:** automated information system; working, development, and operational documentation

**DOI:** 10.14357/08696527140311

### References

1. Zatsarinnyy, A. A., and Y. S. Ionenkov. 2007. Nekotorye aspekty vybora tekhnologii postroeniya informatsionno-telekommunikatsionnykh setey [Some aspects of information-telecommunication system creation technology choice]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 17:5–16.
2. Zatsarinnyy, A. A., and Y. S. Ionenkov. 2009. Metodika vybora tekhnicheskikh sredstv dlya postroeniya telekommunikatsionnykh setey [A technique of rational choice for technical means for telecommunication networks creation]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* Add. Issue:4–14.

3. Zatsarinnyy, A. A., Y. S. Ionenkov, and S. V. Kozlov. 2010. *Nekotorye voprosy proektirovaniya informatsionno-telekommunikacionnykh system* [Some questions of information-telecommunication systems development] Moscow: IPI RAN. 218 p.
4. GOST 2.102-68. 2002. Edinaya sistema konstruktorskoy dokumentatsii. Vidy i kompletnost' konstruktorskikh dokumentov [Unified systems for design documentation. Types and sets of design documentations]. Moscow: StandardInform Publs.
5. GOST 2.601-95. 2004. Edinaya sistema konstruktorskoy dokumentatsii. Ekspluatatsionnye dokumenty [Unified systems for design documentation. Exploitative documents]. Moscow: StandardInform Publs.
6. GOST 2.105-95. 2002. Edinaya sistema konstruktorskoy dokumentatsii. Obshchie trebovaniya k tekstovym dokumentam [Unified systems for design documentation. General requirements for textual documents]. Moscow: StandardInform Publs.
7. GOST 19.101-77. 2010. Edinaya sistema programmnoy dokumentatsii. Vidy programm i programmnykh dokumentov [United system for program documentation. Types of programs and program documents]. Moscow: StandardInform Publs.
8. GOST 34.201-89. 2005. Kompleks standartov na avtomatizirovannye sistemy. Vidy, kompletnost' i oboznachenie dokumentov pri sozdaniii avtomatizirovannykh sistem [Information technology. Set of standards for automated systems. Types, sets and indication of documents for automated systems making]. Moscow: StandardInform Publs.
9. RD 50-34.698-90. 2005. Kompleks standartov na avtomatizirovannye sistemy. Avtomatizirovannye sistemy. Trebovaniya k soderzhaniyu dokumentov [Information technology. Set of standards for automated systems. Requirements to document's content]. Moscow: StandardInform Publs.
10. GOST R 51583-2000. 2003. Zashchita informatsii. Poryadok sozdaniya avtomatizirovannykh sistem v zashchishchennom ispolnenii. Obshchie polozheniya [Data protection. A schedule of secured automated information system creation. Main theses]. Moscow: StateStandard of Russia.
11. GOST RV 15.211-2002. 2004. Sistema razrabotki i postanovki produktsii na proizvodstvo. Voennaya tekhnika. Poryadok razrabotki programm i metodik ispytaniy opytnykh obraztsov izdeliy [System of development and launch new products on production. Military technics. A schedule of program and testing techniques development of product prototypes]. Moscow: StateStandard of Russia.
12. GOST R 21.1703-2000. 2004. Sistema proektnej dokumentatsii dlya stroitel'stva. Pravila vypolneniya rabochey dokumentatsii provodnykh sredstv svyazi [System of design documents for construction. Rules of executing of working documents for wire communication systems]. Moscow: StateConstruct of Russia.

Received July 23, 2014

## Contributors

**Zatsarinnyy Alexander A.** (b. 1951) — Doctor of Science in technology, professor, Deputy Director, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; azatsarinny@ipiran.ru

**Ionenkov Yurij S.** (b. 1956) — senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; uionenkov@ipiran.ru

## МОНИТОРИНГ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ КАК ОБЛАЧНЫЙ СЕРВИС

*O. A. Бондаренко<sup>1</sup>, К. И. Волович<sup>2</sup>, В. А. Кондрашев<sup>3</sup>*

**Аннотация:** Рассматриваются аспекты организации мониторинга информационной безопасности (ИБ) информационных систем (ИС) на основе технологий SIEM (Security information and event management) в качестве сервиса. Сервис представляется центром обработки данных (ЦОД) на основе облачных технологий. Обеспечение качества услуги по мониторингу ИБ в облаке требует определенной модификации традиционной процедуры сбора и обработки данных в SIEM. Предлагается использование специализированных компонентов-посредников сбора первичных данных о функционировании ИС, что позволяет обеспечить сопоставимое качество мониторинга ИБ для локальных и облачных SIEM.

**Ключевые слова:** SIEM; облако; облачный сервис; мониторинг информационной безопасности; виртуализация вычислений

**DOI:** 10.14357/08696527140312

### 1 Введение

Современные концепции обеспечения безопасности ИС предполагают наличие в составе ИС средств противодействия угрозам, а также средств постоянного мониторинга состояния ИБ. Традиционно такие средства внедряются в ИС и функционируют в ней как отдельная подсистема на программно-аппаратных средствах ИС.

Однако в настоящее время в информационных технологиях активно развиваются направления, связанные с виртуализацией вычислений в ЦОД и предоставлении программного обеспечения (ПО) и вычислительных платформ как сервиса. Такие технологии получили название «облачных вычислений» (cloud computing). Задача мониторинга ИБ может рассматриваться как сервис, предоставляемый ЦОД информационным системам, являющимся его клиентами.

Далее в статье проанализированы основные концепции построения систем мониторинга ИБ как компонентов ИС, рассмотрены современные подходы к реализации мониторинга ИБ на основе облачных технологий, выработаны предложения по изменению традиционной архитектуры систем мониторинга ИБ при их применении в качестве облачного сервиса.

---

<sup>1</sup>Институт проблем информатики Российской академии наук, olga@ipi.ac.ru

<sup>2</sup>Институт проблем информатики Российской академии наук, kv@ipi.ac.ru

<sup>3</sup>Институт проблем информатики Российской академии наук, vd@ipi.ac.ru

## **2 Средства автоматизации мониторинга информационной безопасности**

Одним из современных направлений мониторинга ИБ является использование систем SIEM [1]. Основным назначением SIEM является сбор и анализ информации о функционировании ИС от всех программно-аппаратных компонентов, входящих в нее, с целью обнаружения аномалий функционирования, которые могут представлять угрозу ИБ.

Принцип работы SIEM состоит в том, что средства сбора журналов извлекают журналы различных компонентов ИС и заносят их в специализированную базу данных. Дополнительно средства сетевого мониторинга, входящие в SIEM, могут журнализировать в той же базе сетевой трафик ИС. На основе данных, размещенных в базе компонентов, SIEM, называемый коррелятором событий, находит связанные события и строит цепочки корреляции.

Определенные особенности построенных цепочек корреляции (например, повторяемость определенных событий, реакции средств предотвращения атак или межсетевого экранирования на определенный тип сетевого трафика, нетипичная активность пользователей системы) могут служить индикатором того, что ИБ системы находится под угрозой. Для принятия решения об угрозе ИБ SIEM производит сравнение построенных цепочек корреляции событий с шаблонами угроз и принимает решение о потенциальной возможности нарушения ИБ.

Шаблоны угроз разрабатываются поставщиком SIEM и предоставляются как регулярные обновления для SIEM.

В настоящее время ряд крупных производителей ПО предлагает коммерческие реализации SIEM, готовые к внедрению как во вновь создаваемых, так и в существующих ИС.

Лидерами в области SIEM, по данным агентства Gartner, являются [2]:

- IBM QRadar;
- Hewlett Packard Arc Sight;
- Splunk SIEM;
- McAfee SIEM;
- LogRhythm SIEM.

Каждый из указанных производителей имеет собственную лабораторию сетевой безопасности, анализирующую современные методики сетевых атак, действия злоумышленников, уязвимости в ПО известных продуктов. На основании полученных данных лаборатории разрабатывают шаблоны уязвимостей для систем SIEM и предоставляют их пользователям SIEM в качестве обновлений. Причем, поскольку шаблоны учитывают уязвимости в ПО системного и прикладного ПО, SIEM позволяет защитить ИС до того, как производитель соответствующего ПО выпустит обновления.

Несмотря на различия в реализации, предоставляемых сервисах и т. д., общая идея всех систем SIEM — обработать большой объем первичных данных,

не поддающийся анализу человеком, и выдать небольшой набор потенциальных угроз, который может быть проанализирован администратором системы.

Действительно, число записей в системных и прикладных журналах автоматизированной системы, содержащей всего несколько серверов, может доходить до сотен тысяч в день, известны автоматизированные системы крупных корпораций, где количество записей исчисляется миллиардами. Такой объем данных в общем случае не представляет интереса без средств автоматизации их обработки, поэтому идея SIEM не является новинкой, однако промышленные реализации, способные надежно функционировать в условиях реальной эксплуатации, появились только в последние несколько лет. Современные системы SIEM позволяют свести указанный огромный объем данных всего к нескольким (до десятков) сообщениям о потенциальном нарушении ИБ в день, что вполне доступно для анализа и реагирования со стороны администратора ИС.

Внедрение указанных систем SIEM требует развертывания в составе ИС пользователя SIEM следующих программно-аппаратных компонентов:

- СУБД единой базы данных событий ИБ;
- средств сбора и консолидации записей о событиях ИБ в единой базе данных;
- средств анализа записей о событиях ИБ путем построения цепочек корреляции событий;
- средств сопоставления построенных цепочек корреляции с шаблонами уязвимостей с целью обнаружения фактов нарушения ИБ, а также потенциальных уязвимостей;
- средств уведомления администратора и инструментария разбора потенциально опасных ситуаций с целью обнаружения и ликвидации уязвимостей.

Данные компоненты требуют выделения определенных ресурсов (на физических либо виртуальных платформах) в ИС пользователя SIEM и в результате своего функционирования обеспечивают своевременное информирование обслуживающего персонала о потенциальных угрозах ИБ и предоставляют средства разбора и ликвидации инцидентов ИБ.

### **3 Виртуализация сервисов обнаружения инцидентов информационной безопасности**

В настоящее время средства SIEM традиционно функционируют совместно с ИС, для которой осуществляют мониторинг ИБ. Однако можно отметить, что весь описанный выше функционал и средства его реализации могут быть размещены на удаленном ресурсе, имеющем сетевой доступ к компонентам ИС. Поэтому следующим этапом развития SIEM стала его реализация как облачного сервиса (cloud service). Облачные технологии являются на сегодняшний день быстрорастущим сегментом рынка информационных технологий [3, 4]. Облач-

ные технологии основываются на концепции виртуализации вычислений и предоставлении абонентам услуг по обработке, хранению и передаче информации с использованием мощностей ЦОД.

Обычно выделяют три типа ресурсов, предоставляемых облачным ЦОД абоненту в зависимости от его потребностей [5]:

- (1) ПО как сервис (Software as a Service — SaaS);
- (2) платформа как сервис (Platform as a Service — PaaS);
- (3) инфраструктура как сервис (Infrastructure as a Service — IaaS).

В рамках SaaS пользователю предоставляется доступ к определенному ПО (как правило, прикладному), функционирующему на виртуальной инфраструктуре ЦОД. При этом с точки зрения пользователя он получает полный функционал ПО со своими данными, настройками и т. д., аналогичными получаемым при локальной установке ПО, а все мероприятия по эксплуатации аппаратной составляющей ЦОД, обновлениям прикладного ПО, резервному копированию и восстановлению осуществляются силами организации, предоставляющей сервис.

Примером SaaS может служить облачный сервис определения подозрительных транзакций по кредитным картам, используемый многими банками при осуществлении платежей. Данный сервис позволяет банку при инициации клиентом платежа по кредитной карте проверить достоверность запрашиваемого платежа в режиме онлайн простым обращением к соответствующему сервису SaaS через Интернет. При этом банк не размещает у себя никаких дополнительных аппаратно-программных средств и не несет расходов по поддержанию инфраструктуры. Другим примером SaaS является облачный вариант 1С:Бухгалтерии.

В рамках PaaS пользователю предоставляется не конечное прикладное ПО, а комплекс системного ПО, серверов приложений и баз данных, позволяющих развернуть на средствах ЦОД свое собственное прикладное ПО. При этом вопросами эксплуатации системного ПО и серверов приложений занимается персонал ЦОД.

В рамках IaaS пользователю предоставляется виртуальная инфраструктура, позволяющая полностью создавать собственную ИС, включая развертывание операционных систем, серверов приложений, баз данных и прикладного ПО. Пользователю предоставляется полная возможность администрирования и эксплуатации созданного комплекса. Зона ответственности персонала ЦОД в данном случае ограничивается поддержанием работоспособности виртуальной инфраструктуры.

Если рассматривать SIEM как облачный сервис, то очевидно, что он полностью укладывается в концепцию SaaS. Действительно, провайдер сервиса может полностью развернуть на виртуальной инфраструктуре средства SIEM, перечисленные в разд. 2, и предоставить пользователю сервиса собственный экземпляр SIEM, включая базу данных, средства сбора и корреляции событий и прочий инструментарий SIEM.

Такие системы в настоящее время созданы, и провайдеры предоставляют облачные услуги по мониторингу ИБ, называя их иногда “SIEM as a Service”, что в сокращении также звучит как SaaS.

Необходимо отметить, что такие системы действуют по тому же принципу, что и локально установленные SIEM, т. е. собирают журналы от компонентов ИС и передают их в базу данных событий, расположенную в облаке. Сбор событий может осуществляться как специальными агентами, установленными в ИС, так и средствами самого сервиса, имеющими доступ к компонентам системы.

#### **4 Структура облачного сервиса обнаружения инцидентов информационной безопасности**

Описанная в разд. 3 схема работы облачного сервиса обнаружения инцидентов ИБ имеет очевидный недостаток, связанный с необходимостью передавать весь объем первичных данных (событий) из ИС в облачный сервис. Другим недостатком данной структуры является необходимость открывать доступ облачному сервису к большому количеству компонентов ИС. В случае с локальным расположением SIEM и ИС данный аспект не является помехой для использования. Кроме того, при локальном расположении имеется возможность использования в анализе данных о внутреннем сетевом трафике, что невозможно при использовании SaaS.

Для исключения указанного недостатка при использовании SaaS предлагаются модифицировать состав данных, передаваемых от ИС в облачный SIEM. Для этого предлагается использовать компонент-посредник SIEM (LOG-Proxy), внедряемый в ИС и осуществляющий сбор и предварительную обработку информации о событиях, включая записи в системных и прикладных журналах, а также (при необходимости) данные о сетевом трафике. Посредник может размещаться как на дополнительно выделяемых аппаратно-программных (или виртуальных) ресурсах ИС, так и на существующих мощностях.

Предварительная обработка состоит в том, что на основе первичных данных LOG-Proxy должен формировать информационные объекты, интегрально описывающие состояние ИБ. Такие «отпечатки» текущего состояния журналов и сетевого трафика, занимающие существенно меньший объем, чем первичные данные, должны передаваться в облако для последующего корреляционного анализа событий и выявления признаков нарушения ИБ. Объекты должны формироваться LOG-Proxy на основе правил, размещаемых на разделяемом ресурсе облачного сервиса SIEM и регулярно обновляемых производителем SIEM вместе с шаблонами уязвимостей.

#### **5 Заключение**

Для адаптации существующих средств SIEM к условиям функционирования в качестве облачного сервиса предлагается изменить способ сбора и предоставле-

ния для анализа первичных данных путем внедрения в ИС компонента-посредника LOG-Proxy, обеспечивающего предварительную обработку данных перед передачей их в SIEM.

Предлагаемый подход позволяет повысить эффективность обнаружения уязвимостей облачным сервисом за счет использования данных о сетевом трафике, недоступных при прямом переносе функционала SIEM в облако. Кроме того, существенно снижается нагрузка и требования к каналам связи между SIEM-провайдером и пользователем за счет уменьшения объема трафика, передаваемого между ними. Относительным недостатком такого подхода можно считать необходимость разработчику SIEM создавать и поддерживать «облачную» версию своего продукта, включающую LOG-Proxy, а также измененную базу данных и средства корреляции событий. Однако производители обычно идут на выпуск специальной версии ПО, ориентированного на применение в облачных сервисах, поэтому задача такой доработки не выходит за рамки обычной практики разработки ПО SaaS на базе существующих локальных продуктов.

## **Литература**

1. *Antonopoulos N., Gillam L.* Cloud computing: Principles, systems and applications. — London: Springer, 2010. 271 p.
2. *Kavanagh K., Nicolett M., Rochford O.* Magic quadrant for security information and event management // Gartner, June 25, 2014. <http://www.alienvault.com/marketing/siem-magic-quadrant-2014>.
3. Облачные сервисы (рынок России) // TAdviser, 25.08.2013. [http://www.tadviser.ru/index.php/Статья:Облачные\\_сервисы\\_\(рынок\\_России\)](http://www.tadviser.ru/index.php/Статья:Облачные_сервисы_(рынок_России)).
4. *Prokhorov A.* Russia cloud services market 2012 analysis and 2013–2017 forecast // IDC, September 2013. <http://www.idc.com/getdoc.jsp?containerId=ES33V>.
5. *Miller D. R.* Security information and event management (SIEM) implementation. — New York: McGraw-Hill, 2010.

*Поступила в редакцию 25.08.14*

---

## **MONITORING OF INFORMATION SECURITY AS A CLOUD SERVICE**

*O. Bondarenko, K. Volovich, and V. Kondrashev*

Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str.,  
Moscow 119333, Russian Federation

**Abstract:** The article considers aspects and issues of information security monitoring through systems SIEM (Security information and event management) as a service. Service is provided by the cloud data center. Quality assurance monitor-

ing services in the cloud requires some modification of the traditional procedure of collecting and processing data in SIEM. Using specialized mediations primary data collection allows receiving comparable quality of monitoring of information security between local and clouding SIEM.

**Keywords:** SIEM; cloud; cloud service; monitoring of information security; virtualization

**DOI:** 10.14357/08696527140312

## References

1. Antonopoulos, N., and L. Gillam. 2010. *Cloud computing: Principles, systems and applications*. London: Springer. 271 p.
2. Kavanagh, K., M. Nicolett, and O. Rochford. June 25, 2014. Magic quadrant for security information and event management. Gartner Inc. Available at: <http://www.alienvault.com/marketing/siem-magic-quadrant-2014> (accessed August 25, 2014).
3. Oblachnye servisy (rynokRossii) [Cloud services (Russian market)]. 25.08.2013. TAdviser. Available at: [http://www.tadviser.ru/index.php/Статья:Облачные\\_сервисы\\_\(рынок\\_России\)](http://www.tadviser.ru/index.php/Статья:Облачные_сервисы_(рынок_России)) (accessed August 25, 2014).
4. Prokhorov, A. 2013. Russia cloud services market 2012 analysis and 2013–2017 forecast. IDC. September 2013. Available at: <http://www.idc.com/getdoc.jsp?containerId=ES33V> (accessed August 25, 2014).
5. Miller, D. R. 2010. *Security information and event management (SIEM) implementation*. New York: McGraw-Hill. 465 p.

*Received August 25, 2014*

## Contributors

**Bondarenko Olga A.** (b. 1970) — junior scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; olga@ipi.ac.ru

**Volovich Konstantin I.** (b. 1970) — senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; kv@ipi.ac.ru

**Kondrashev Vadim A.** (b. 1963) — senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vd@ipi.ac.ru

## АВТОМАТИЗИРОВАННАЯ ИНФОРМАЦИОННАЯ СИСТЕМА КОМПЛЕКСНОГО МОНИТОРИНГА ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ\*

*А. А. Сорокин<sup>1</sup>, А. Г. Тарасов<sup>2</sup>, С. П. Королёв<sup>3</sup>*

**Аннотация:** Рассматриваются вопросы создания информационной системы, предназначенной для оценки состояния сети передачи данных с использованием протоколов SNMP (Simple Network Management Protocol) и NetFlow. Предложена концепция системы мониторинга для корпоративной сети передачи данных, построенной на основе топологии «звезда» и определены базовые требования к ней. Представлено описание программной платформы, разработанной на основе открытых стандартов и технологий и состоящей из набора специализированных компьютерных модулей, интегрированных с системой Zabbix. Они позволяют в рамках единого пользовательского интерфейса проводить комплексный анализ состояния сети передачи данных. Описана система архивации инструментальных данных и результатов их обработки, обеспечивающая максимальную скорость обработки данных как в режиме оперативных наблюдений, так и ретроспективного анализа. Приведен пример апробации созданных программных средств на основе Региональной компьютерной сети (РКС) ДВО РАН.

**Ключевые слова:** корпоративная сеть передачи данных; информационная система; сетевой мониторинг; SNMP; NetFlow; Zabbix; протокол; система управления базами данных

**DOI:** 10.14357/08696527140313

### 1 Введение

С развитием телекоммуникационных сетей, информационных сервисов, а также лавинообразным ростом числа их пользователей у операторов связи возникает проблема обеспечения качества предоставляемых услуг. Она решается за счет

\* Работа выполнена при поддержке РФФИ (проект 12-07-31149 мол\_а), Комплексной программы фундаментальных исследований Дальневосточного отделения РАН «Дальний Восток», грантов ДВО РАН 12-І-П14-02 и 12-ІІІ-А-01И-013.

<sup>1</sup> Вычислительный центр Дальневосточного отделения Российской академии наук,  
alsor@febras.net

<sup>2</sup> Вычислительный центр Дальневосточного отделения Российской академии наук,  
taleks@as.khb.ru

<sup>3</sup> Вычислительный центр Дальневосточного отделения Российской академии наук,  
serejk@febras.net

внедрения систем мониторинга и управления сетью, которые должны информировать администраторов сети о проблемных ситуациях (событиях), возникающих вследствие ошибок средств связи, использования прикладных информационных сервисов, генерирующих большое число соединений, проблем в области информационной безопасности и т. п. Указанные системы должны обладать инструментами для проведения качественного и количественного анализа трафика сети, оценки состояния компонентов информационно-телеинформационной инфраструктуры.

Выявление аномальных событий в работе сети позволяет своевременно устранить или минимизировать проблему и тем самым обеспечить надлежащий уровень сервиса как для работы конечных потребителей, использующих стандартные пользовательские приложения, так и информационных сервисов (GRID, системы видео-конференц-связи и др.), для работы которых требуются ресурсы с гарантированной пропускной способностью. Решение вышеперечисленных задач связано с проведением не только технических, но и исследовательских работ, таких как анализ потоков данных, разработка или адаптация алгоритмов их обработки и т. п. При этом важным исходным материалом для выполнения этих работ является инструментальная информация, получаемая в результате оценки работы сети с использованием специализированных сетевых протоколов, среди которых наиболее востребованными являются SNMP [1, 2], NetFlow [3], ICMP (Internet Control Message Protocol) [4]. Первый представляет собой иерархическую систему различных счетчиков и средства управления устройствами, второй — протокол для учета качественных и количественных показателей трафика сети. Протокол ICMP используется, в частности, для контроля доступности хостов в IP-сети. На базе этих и других протоколов работают многочисленные системы мониторинга информационно-телеинформационной инфраструктуры (Ganglia, Nagios, Zabbix, Netams и др.). Большинство из них помимо инструментов учета и визуализации состояния сети позволяют создавать шаблоны поведения и реакции на возникающие события.

Как правило, системы мониторинга достаточно эффективно решают задачи сбора и обработки данных в своей узкоспециализированной области. Например, Ganglia [5] ориентирована на работу в сети вычислительного кластера. Поэтому, как правило, такие системы не обладают возможностью взаимодействия друг с другом на уровне сервисов, что исключает взаимное использование различных архивов данных и средств их интерпретации. Таким образом, для поддержки и комплексной оценки функционирования сети и ее ресурсов необходимо поддерживать и работать с множеством различных специализированных информационных систем. Это создает очевидные ограничения и неудобства, которые приводят к снижению скорости реакции на возникающие инциденты и увеличению вероятности возникновения ошибки администратора при работе с большим числом программных инструментов.

Несмотря на многообразие информационных систем, до настоящего времени нет открытых решений, которые бы интегрировали на одной программной плат-

форме инструменты для работы с протоколами SNMP и NetFlow с возможностью централизованной обработки данных, получаемых от средств управления сетью. Совместный анализ данных как минимум этих двух протоколов позволит получить качественные и количественные показатели функционирования сети и более эффективно выявлять случаи воздействия на нее.

В Вычислительном центре (ВЦ) ДВО РАН ведутся работы по созданию прототипа автоматизированной информационной системы комплексного мониторинга сети (далее — АИСКМС), предназначенный для решения задач мониторинга телекоммуникационной сети с использованием протоколов SNMP (версии 2) и NetFlow (nonsampled, версии 5) [6].

## **2 Концепция комплексного мониторинга сети**

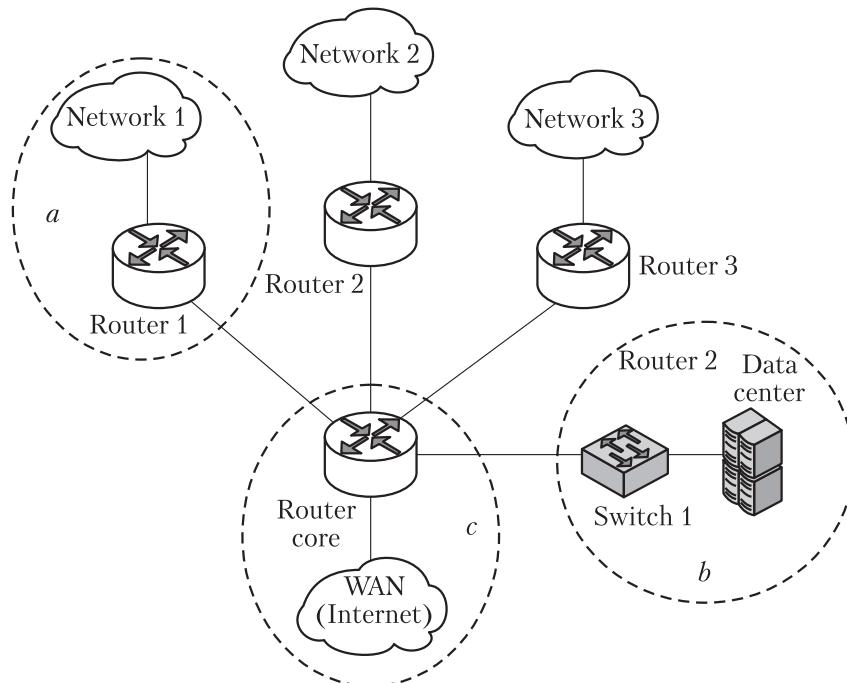
Разные виды телекоммуникационных сетей имеют свои цели, задачи и назначение. Исходя из этого, при их создании используются те или иные технологии передачи данных, обеспечения информационной безопасности, определяются требования к надежности и срокам реагирования на возникающие инциденты. В связи с этим система мониторинга должна учитывать особенности сети, для которой она создается.

Разрабатываемая АИСКМС ориентирована на работу в корпоративных сетях, построенных на основе топологии «звезда» со структурой управления, отраженной на рис. 1.

Такая сеть объединяет следующие условные базовые элементы:

- (1) сегмент сети (*a*), состоящий из локальных сетей организаций или групп пользователей, находящихся в отдельных регионах или выделенных для производственной необходимости;
- (2) центр данных (*b*), обеспечивающий поддержку централизованных системных и прикладных информационных сервисов и систем сети (DNS (domain name system), www (world-wide web), mail и т. п.);
- (3) периметр (*c*), решающий задачи интеграции разных сегментов и их доступа в глобальные сети.

Оценка и контроль работы этих частей строится на основе данных о состоянии оборудования связи (интерфейсы, загрузка процессора, использование памяти, работа инженерных блоков, состояние окружающей среды и т. п.), а также характеристиках потоков данных между ними (объемы передаваемой информации, виды протоколов, сетевые соединения и т. п.). При использовании одинакового набора исходных инструментальных данных каждая часть сети, исходя из решаемых задач, требует формирования отдельных правил реагирования, составленных из определенного числа и видов параметров и времени реакции на них.

**Рис. 1** Типовая схема корпоративной сети

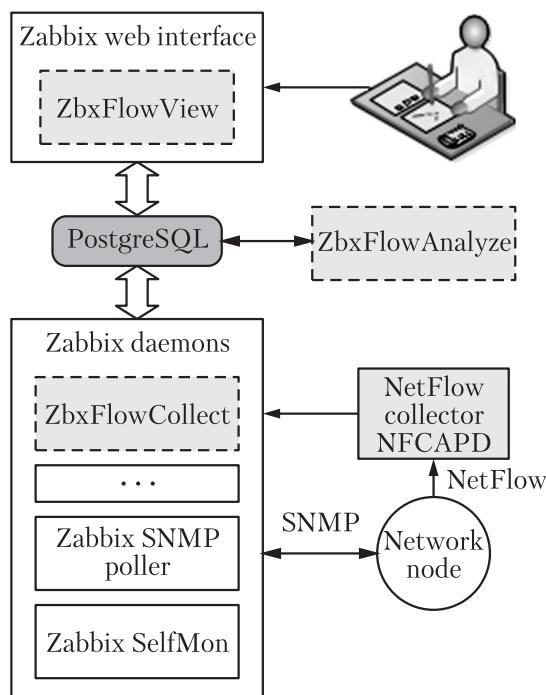
Для решения этих задач АИСКМС должна обеспечить:

- сбор инструментальных данных с управляющих узлов сети по различным протоколам и передачу их в реальном времени в специализированный коллектор или базу данных;
- формирование структурированных архивов данных с первичной информацией и результатами ее обработки, обеспечение их сохранности и доступности;
- предоставление инструментов для обработки и совместного анализа разных инструментальных данных с целью получения оценки состояния как сети в целом, так и отдельных ее частей с возможностью изменения числа параметров, по которым будет производиться такой анализ;
- средства визуализации данных и построения карт сети;
- оповещение заинтересованных лиц о возникающих событиях по заранее заданным шаблонам и правилам (нагрузка, скорость передачи данных, превышение критического числа ошибок по определенным параметрам и т. п.);
- наличие масштабируемой программной архитектуры, позволяющей добавлять в информационную систему поддержку новых протоколов мониторинга и инструментов для работы с инструментальными данными.

В качестве ядра платформы, создаваемой АИСКМС, использован свободно распространяемый (лицензия GPL v2) программный продукт Zabbix<sup>1</sup> [7]. Он предоставляет средства мониторинга сети по протоколам SNMP и IPMI (Intelligent Platform Management Interface), а также интерфейс программирования приложений (API) для разработки виджетов визуализации, доступа к архивам данных и инструментам для работы с ними. На этой основе авторами разработан набор компьютерных модулей, которые реализуют технологии работы с данными NetFlow и обеспечивают их интеграцию с ресурсами Zabbix, образуя единую информационную среду.

### **3 Архитектура автоматизированной информационной системы комплексного мониторинга сети**

Автоматизированная информационная система комплексного мониторинга сети представляет собой комплекс специализированных компьютерных программ (модулей), взаимодействующих друг с другом с использованием архитектуры клиент–сервер (рис. 2).



**Рис. 2** Схема взаимодействия модулей АИСКМС

---

<sup>1</sup><http://www.zabbix.com>.

### 3.1 Модуль сбора исходных инструментальных данных мониторинга

Система предназначена для работы с двумя базовыми протоколами мониторинга — SNMP и NetFlow. На устройствах сети (network node) производятся штатные настройки, связанные с обеспечением поддержки указанных протоколов, заданием параметров поставки данных и правил безопасности.

Данные SNMP собираются с устройств сети, обрабатываются и помещаются в архив системы, реализованный на основе системы управления базами данных (СУБД) PostgreSQL, штатными средствами Zabbix.

Для приема и хранения данных NetFlow организован коллектор на основе пакета `nfcapd`<sup>1</sup>. Для выполнения операций с записями разработан программный модуль `ZbxFlowCollect`. Он управляет мастер-процессом Zabbix и решает задачи проверки исходных данных и их передачи из коллектора NetFlow в СУБД PostgreSQL. Схема данных хранения записей NetFlow представлена на рис. 3.

Для межпроцессного взаимодействия, позволяющего `ZbxFlowCollect` получать наборы данных NetFlow от `nfcapd`, используется программа компонента, основанная на unix pipe [8].

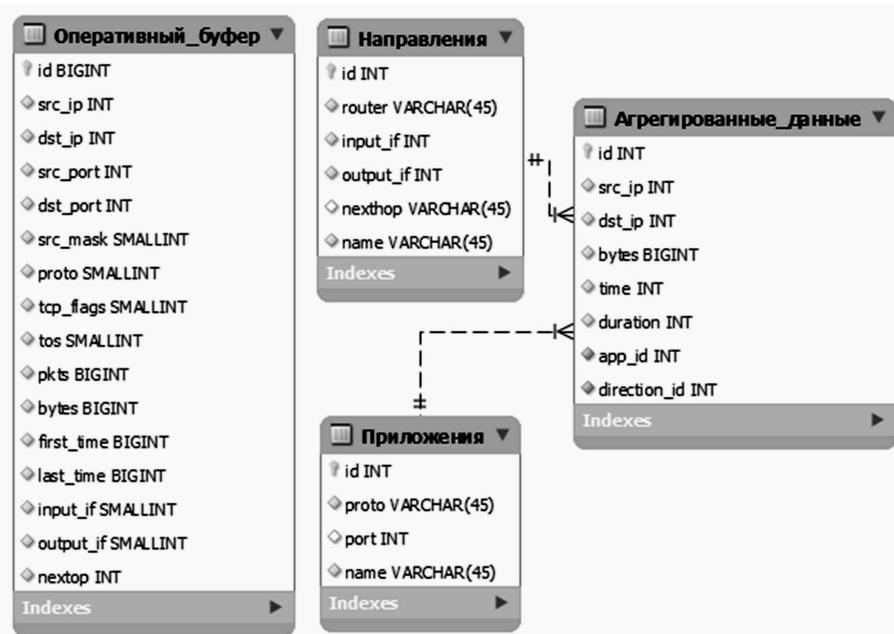


Рис. 3 Схема базы данных PostgreSQL для хранения записей АИСКМС

<sup>1</sup><http://nfdump.sourceforge.net>.

### **3.2 Модуль анализа данных мониторинга**

Для анализа данных NetFlow разработан модуль ZbxFlowAnalyze. Он позволяет создавать шаблоны для автоматической экспресс-оценки данных мониторинга с целью выявления различных событий, таких как вирусная активность, сканирование сервисов и устройств Сети, DDoS (Distributed Denial of Service) атаки на периметр и т. п. На основе заданных правил модуль производит анализ заголовков пакетов четвертого уровня модели OSI (open systems interconnection) [9], находящихся в оперативном буфере системы.

При обнаружении совпадения генерируется соответствующее уведомление с использованием средств платформы Zabbix (рис. 4).



**Рис. 4** Схема отправки уведомления о событии модуля ZbxFlowAnalyze

В модуле помимо использования существующих алгоритмов и методик работы с данными NetFlow [10, 11] заложена возможность подключения дополнительных систем, позволяющих решать узкоспециализированные задачи мониторинга.

### **3.3 Модуль ZbxFlowView**

Для работы с созданными модулями реализован набор веб-интерфейсов в виде единого модуля ZbxFlowView. Он позволяет производить настройку АИСКМС, а также работу с инструментальными данными и результатами их обработки.

Интерфейс выполнен как часть единого веб-интерфейса Zabbix на основе его программных классов и виджетов (CWidget, CTag) и функционирует в рамках общей политики безопасности, т. е. единой системы ролей пользователей и прав доступа к данным. В качестве источника информации для генерации страниц выступает СУБД PostgreSQL, куда поступает вся информация от модулей АИСКМС.

## **4 Система архивации данных автоматизированной информационной системы комплексного мониторинга сети**

Работа с данными реального времени, хранение значений различных показателей и статистики требуют соответствующих вычислительных ресурсов и систем хранения. Поэтому при настройке системы мониторинга важно определить перечень используемых параметров по всем протоколам и режимы работы системы, т. е. когда система будет оперировать данными реального времени, а когда —

результатами обработки их значений. Объем хранимой в базе данных информации, а именно значений различных показателей мониторинга, напрямую зависит от топологии и характеристик сети, а также сетевой активности.

Особенностью работы с SNMP-данными является то, что все значения показателей, поставляемые с узлов сети, сразу помещаются в базу данных. Для каждого устройства в Zabbix настраивается интервал опроса и длительность хранения полученных данных. Поскольку число наблюдаемых показателей конечно, то при фиксированном интервале наблюдений и требуемой длительности хранения значений отсутствуют сложности в хранении и обработке таких массивов данных. На данный момент АИСКМС собирает 1095 SNMP-показателей, значения которых хранит в течение одного года.

Исходные записи NetFlow (датаграммы) имеют более сложную структуру. В отличие от SNMP-показателей, количество которых фиксировано, Netflow-записей с качественными и количественными данными по сетевым соединениям может быть неограниченное количество.

Для работы с исходными инструментальными данными NetFlow в АИСКМС были предложены следующие подходы:

- (1) АИСКМС в первую очередь ориентирована на работу с nonsampled версией NetFlow, позволяющей непрерывно получать весь объем оперативной информации с сенсоров сети. На текущем этапе развития системы для хранения и обработки был выделен набор полей датаграмм NetFlow, представленный в таблице;
- (2) при высокой сетевой активности сенсоры сети могут генерировать существенный объем инструментальной информации NetFlow (миллионы записей в час). Поэтому для ее эффективной обработки и хранения были разработаны правила агрегации и архивирования сформированного (см. п. 1) набора NetFlow-данных:

Набор полей NetFlow v5, используемых в АИСКМС

№	Название поля	Описание поля
1	srcaddr	IP-адрес отправителя
2	dstaddr	IP-адрес получателя
3	srcport	Порт отправителя
4	dstport	Порт получателя
5	proto	Номер IP-протокола по классификации IANA
6	tcp_flags	Логическая сумма флагов TCP/IP-пакета
7	dPkts	Число переданных IP-пакетов
8	dOctets	Число переданных байт
9	input	SNMP-индекс входящего интерфейса маршрутизатора
10	output	SNMP-индекс исходящего интерфейса маршрутизатора
11	src_mask	Маска адреса отправителя в CIDR-нотации
12	dst_mask	Маска адреса получателя в CIDR-нотации
13	nexthop	IP-адрес маршрутизатора, выбранного для пересылки пакетов

- (а) в течение  $X$  часов данные сохраняются в исходном виде без агрегирования. Они формируют так называемый оперативный буфер (см. рис. 3), позволяющий проводить максимально подробную оценку текущего состояния сети;
- (б) в течение последующих  $Y$  суток данные агрегируются с интервалом в 5 мин, удаляются поля `tcp_flags`, `src_mask`, `dst_mask`. Поля `input`, `output`, `nexthop`, а также IP-адрес сетевого сенсора наполняют созданный справочник направлений трафика в сети;
- (в) в течение следующих  $Z$  суток данные агрегируются с интервалом в 1 ч. Удаляется поле `dPkts`, а на основе полей `proto`, `srcport` и `dstport` формируется справочник приложений сети. Например, порт 80 по протоколу TCP представляет HTTP-приложение.

Задание временных интервалов ( $X, Y, Z$ ) производится средствами пользовательского интерфейса `ZbxFlowView`. Конкретные значения этих интервалов зависят от задач системы мониторинга, а также от возможностей вычислительной инфраструктуры центра данных, где разворачивается АИСКМС и будет обрабатываться и храниться информация.

## **5 Отличительные особенности автоматизированной информационной системы комплексного мониторинга сети**

Автоматизированная информационная система комплексного мониторинга сети, созданная на основе открытой платформы `Zabbix`, обладает рядом преимуществ:

- (1) лицензия не запрещает ее бесплатное использование, что является важным фактором для небольших корпоративных сетей;
- (2) помимо основных протоколов `SNMP` и `NetFlow` система также поддерживает мониторинг находящихся в сети серверов (потребителей сетевого трафика) с помощью агентов мониторинга и технологии `IPMI`. Полученная информация позволяет проводить расширенный анализ возникающих в сети аномалий;
- (3) в силу открытости базовых компонент АИСКМС может быть расширена, в том числе уже существующими модулями `ZABBIX`, например, для мониторинга платформ виртуализации и облачных сред.

По мнению авторов, перечисленные особенности позволяют рассматривать созданный инструмент как более оптимальное решение для средних по размеру корпоративных сетей по сравнению с существующими коммерческими аналогами, например `Cisco Prime`. Такой вывод обусловлен анализом условий лицензирования и адаптации, присущих большинству коммерческих продуктов. Среди них:

- высокая цена на приобретение и обновление системы и ее компонент;
- закрытость и невозможность развития программного обеспечения для решения специализированных задач;
- зачастую ориентированность на аппаратную платформу одного производителя.

Указанные особенности коммерческих систем не являются их недостатками. Однако при создании комплексных систем мониторинга для сетей с использованием неоднородного аппаратного обеспечения, наличием большого числа поддерживаемых протоколов и технологий мониторинга, необходимостью реализации дополнительного функционала использование открытых систем, таких как АИСКМС, является более оправданным решением.

## **6 Апробация автоматизированной информационной системы комплексного мониторинга сети**

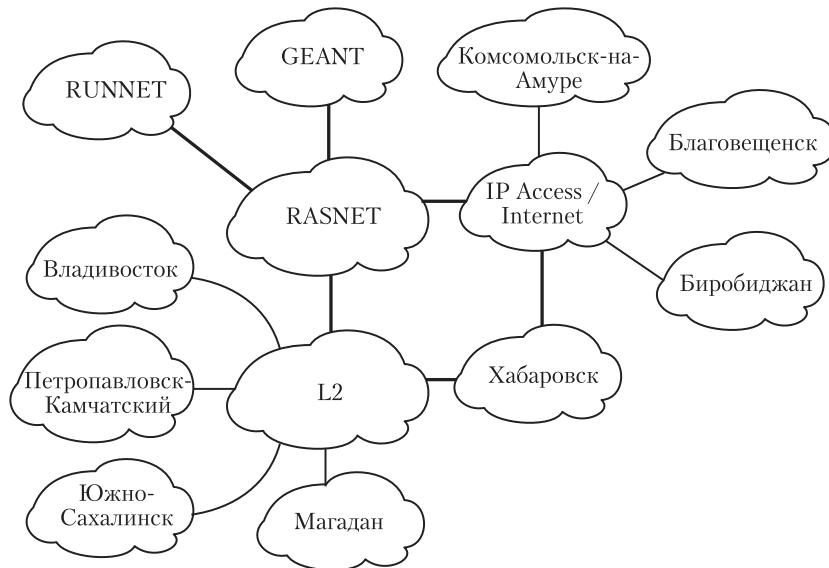
Проверка работы АИСКМС проведена на основе РКС ДВО РАН [12].

Дальневосточное отделение РАН объединяет учреждения, проводящие фундаментальные и прикладные научные исследования по важнейшим проблемам естественных, технических, общественных и гуманитарных наук на территории Дальневосточного федерального округа. Региональная компьютерная сеть имеет следующую архитектуру и структуру управления. Во всех региональных научных центрах ДВО РАН (Хабаровск, Владивосток, Благовещенск, Петропавловск-Камчатский, Магадан, Южно-Сахалинск) созданы и функционируют сегменты и опорные узлы Сети, которые обеспечивают маршрутизацию локальных сетей и работу базовых сетевых сервисов (DNS, mail, www, proxy и т. п.), а также специализированных информационных систем в интересах учреждений ДВО РАН [13].

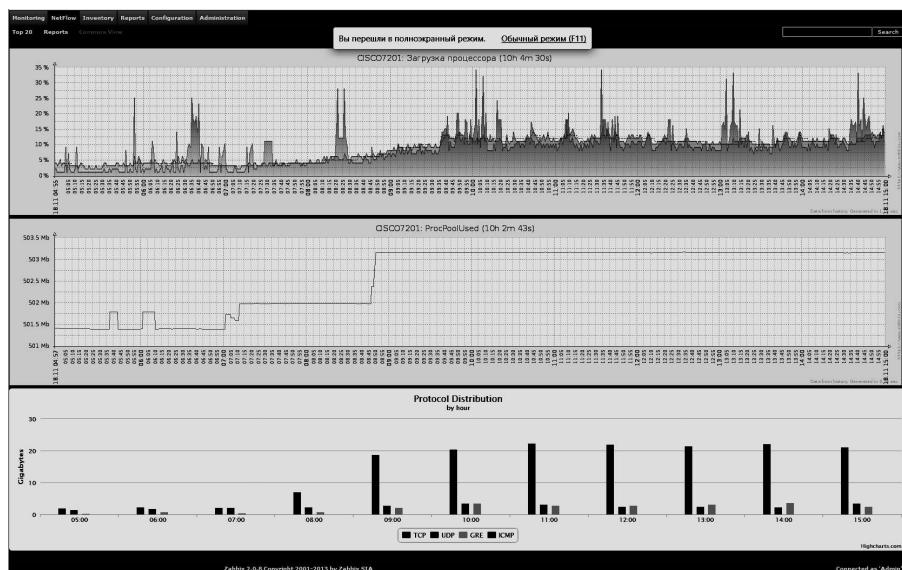
В ВЦ ДВО РАН в г. Хабаровск организован Центр управления РКС ДВО РАН, который обеспечивает настройку и сопровождение базовой системы информационной безопасности, управление службой маршрутизации и качеством услуг связи QoS (quality of service), а также решает задачи по поддержке управляющих систем централизованных информационных сервисов сети.

Для интеграции сегментов Сети и доступа к ресурсам глобальных сетей используются различные виды каналов связи и телекоммуникационных технологий (L2, IP Access и т. п.) (рис. 5).

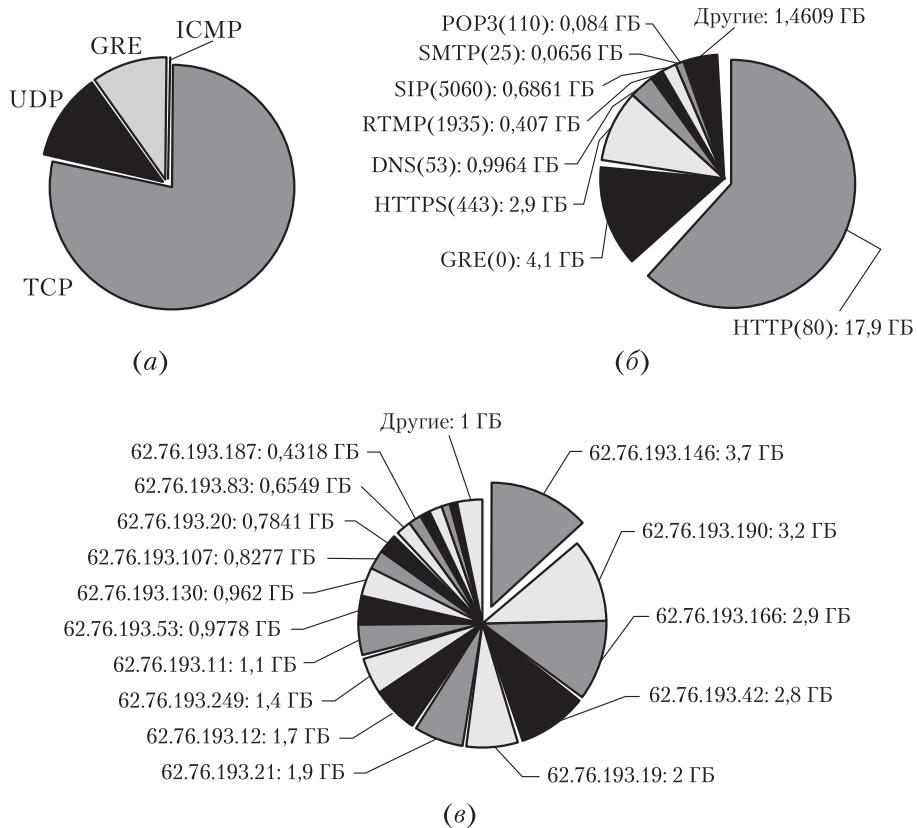
Предварительная работа по установке системы мониторинга Zabbix позволила с использованием данных SNMP создать карту РКС ДВО РАН, включив в нее все маршрутизаторы и коммутаторы ядра сети. Возможность совместной работы с NetFlow позволила в рамках единой информационной среды дополнить информацию о состоянии оборудования и каналов связи сети данными по сетевым соединениям и трафику (рис. 6 и 7).



**Рис. 5** Общая схема каналов связи РКС ДВО РАН



**Рис. 6** Рост нагрузки на пограничный маршрутизатор вместе с ростом потребления трафика



**Рис. 7** Структура трафика пограничного маршрутизатора: (а) по протоколам; (б) по приложениям; (в) по IP-адресам

Это дало возможность быстрее проводить поиск неисправностей в сети и принимать меры по их устранению. Например, анализируя плохую работу выделенного канала связи (потери при передаче, чрезмерная загрузка и т. п.), можно проверить отчеты по ошибкам на интерфейсах, нагрузку маршрутизатора сети (данные SNMP), а также структуру трафика (данные NetFlow).

После локализации проблемы она может быть либо решена, либо передана в виде заявки на обслуживание в организацию, в зоне ответственности которой она находится.

## 7 Заключение

В настоящее время создан действующий прототип АИСКМС, позволяющий решать задачи оценки функционирования телекоммуникационных сетей на основе

совместного анализа данных протоколов SNMP и NetFlow. Это позволяет в рамках единой информационной платформы Zabbix реализовать эффективные инструменты для контроля аппаратной инфраструктуры сетей и их качественных и количественных параметров функционирования.

Разработанные программные компоненты апробированы и внедрены в работу РКС ДВО РАН. Дальнейшее развитие АИСКМС связано с созданием специализированных систем для оценки состояния телекоммуникационной инфраструктуры и аппаратных компонент сетей сейсмологических и деформационных наблюдений [14], распределенных и облачных вычислений ДВО РАН [15].

## **Литература**

1. Hunt R. Protocols and architecture for managing TCP/IP network infrastructures // Comput. Commun., 1997. Vol. 20. No. 2. P. 73–88.
2. Mauro D., Schmidt K. Essential SNMP. — 2nd ed. — Cambridge, MA: O'Reilly Media, 2005. 462 p.
3. Bin L., Chuang L., Jian Q., Jianping H., Ungsunan P. A NetFlow based flow analysis and monitoring system in enterprise networks // Comput. Netw., 2008. Vol. 52. No. 5. P. 1074–1092.
4. Cohen F. Internet holes. Part 1: Internet Control Message Protocol // Netw. Security, 1995. Vol. 1995. Iss. 8. P. 8–10.
5. Massie M., Chun B., Culler D. The Ganglia distributed monitoring system: Design, implementation, and experience // Parallel Comput., 2004. Vol. 30. No. 7. P. 817–840.
6. NetFlow export datagram format. [http://www.cisco.com/en/US/docs/net\\_mgmt/netflow\\_collection\\_engine/3.6/user/guide/format.html](http://www.cisco.com/en/US/docs/net_mgmt/netflow_collection_engine/3.6/user/guide/format.html).
7. Olups R. Zabbix 1.8 network monitoring. — Birmingham, U.K.: Packt Publishing, 2010. 428 p.
8. Jones T. GNU/Linux application programming. — 2nd ed. — Programming ser. Boston, MA: Course Technology, 2008. 486 p.
9. Burkhardt H. J., Krecioch T. H. Standards for the ISO — Reference model for open systems interconnection // Microproc. Microprog., 1982. Vol. 9. Iss. 4. P. 237–244.
10. Ivanovic I. Recommendations for network traffic analysis using the NetFlow protocol. — AMRES NMS Group, 2011. 34 p.
11. Detecting and analyzing network threats with NetFlow. Cisco Systems, 2009. <http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/netflow/configuration/12-4t/nf-12-4t-book/nf-detct-analy-thrts.html>.
12. Khanchuk A. I., Smagin S. I., Sorokin A. A., Makogonov S. V. Regional network of the Far Eastern Branch of RAS // 1st Russian and Pacific Conference on Computer Technology and Applications. — Vladivostok: Far Eastern Branch of the Russian Academy of Sciences, 2010. P. 233–234.
13. Ханчук А. И., Сорокин А. А., Смагин С. И., Королев С. П., Макогонов С. В., Тарасов А. Г., Шестаков Н. В. О развитии информационно-телекоммуникационных систем в Дальневосточном отделении РАН // Информационные технологии и вычислительные системы, 2013. № 4. С. 45–57.

14. Ханчук А. И., Коновалов А. В., Сорокин А. А., Королев С. П., Гаврилов А. В., Бормотов В. А., Серов М. А. Инstrumentальное и информационно-технологическое обеспечение сейсмологических наблюдений на Дальнем Востоке России // Вестник ДВО РАН, 2011. Т. 157. № 3. С. 127–137.
15. Сорокин А. А., Королев С. П., Смагин С. И., Поляков А. Н. Макет отказоустойчивой информационной системы для облачного хранения наборов научных данных // Вычислительные технологии, 2013. Т. 18. № 1. С. 87–95.

Поступила в редакцию 02.12.13

---

---

## AUTOMATED INFORMATION SYSTEM FOR TELECOMMUNICATION NETWORK COMPLEX MONITORING

*A. Sorokin, A. Tarasov, and S. Korolev*

Computer Center of the Far Eastern Branch of the Russian Academy of Sciences,  
65 Kim U Chena Str., Khabarovsk 680000, Russian Federation

**Abstract:** The development of an information system for monitoring the status of a star topology corporate network is considered. Base requirements of the system are specified. The system uses the SMNP (Simple Network Management Protocol) and NetFlow protocols. The system software platform is developed using open standards and technologies. It consists of a set of specialized computer modules integrated with the Zabbix system. The modules provide comprehensive analysis of the network status data under a single user interface. A raw and processed data archiving system is described that provides for maximum processing speed in both operational observations and retrospective data analysis. An example of using the developed program tools is given based on the Regional Computer Network of FEB RAS.

**Keywords:** corporate telecommunication network; information system; network monitoring; SNMP; NetFlow; Zabbix; protocol; database management

**DOI:** 10.14357/08696527140313

### Acknowledgments

The research was supported by the Russian Foundation for Basic Research (project 12-07-31149 mol\_a), Complex Program of Basic Research of FEB RAS “Far East,” and FEB RAS grants 12-I-P14-02 and 12-III-A-011-013.

### References

1. Hunt, R. 1997. Protocols and architecture for managing TCP/IP network infrastructures. *Comput. Commun.* 20(2):73–88.

2. Mauro, D., and K. Shmidt. 2012. *Essential SNMP*. 2nd ed. Cambridge, MA: O'Reilly Media. 462 p.
3. Bin, L., L. Chuang, Q. Jian, H. Jianping, and P. Ungsunan. 2008. A NetFlow based flow analysis and monitoring system in enterprise networks. *Comput. Netw.* 52(5):1074–1092.
4. Cohen, F. 1995. Internet holes. Part 1: Internet Control Message Protocol. *Netw. Security* 1995(8):8–10.
5. Massie, M., B. Chun, and D. Culler. 2004. The Ganglia distributed monitoring system: Design, implementation, and experience. *Parallel Comput.* 30(7):817–840.
6. NetFlow export datagram format. Available at: [http://www.cisco.com/en/US/docs/net\\_mgmt/netflow\\_collection\\_engine/3.6/user/guide/format.html](http://www.cisco.com/en/US/docs/net_mgmt/netflow_collection_engine/3.6/user/guide/format.html) (accessed July 19, 2012).
7. Olups, R. 2010. *Zabbix 1.8 network monitoring*. Birmingham, U.K.: Packt Publishing. 428 p.
8. Jones, T. 2008. *GNU/Linux application programming*. 2nd ed. Programming ser. Boston, MA: Course Technology. 486 p.
9. Burkhardt, H. J., and T. H. Krecioch. 1982. Standards for the ISO—reference model for open systems interconnection. *Microproc. Microprog.* 9(4):237–244.
10. Ivanovic, I. 2011. *Recommendations for network traffic analysis using the NetFlow protocol*. AMRES NMS Group. 34 p.
11. Cisco Systems. 2009. *Detecting and analyzing network threats with NetFlow*. 60 p. Available at: <http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/netflow/configuration/12-4t/nf-12-4t-book/nf-detct-analy-thrts.html> (accessed September 29, 2014).
12. Khanchuk, A. I., S. I. Smagin, A. A. Sorokin, and S. V. Makogonov. 2010. Regional network of the Far Eastern Branch of RAS. *1st Russian and Pacific Conference on Computer Technology and Applications*. Vladivostok: Far Eastern Branch of the Russian Academy of Sciences. 233–234.
13. Khanchuk, A. I., A. A. Sorokin, S. I. Smagin, S. P. Koroljov, S. V. Makogonov, A. G. Tarasov, and N. V. Shestakov. 2013. O razvitiyu informatsionno-telekommunikatsionnykh sistem v Dal'nevostochnom otdelenii RAN [Development of information and telecommunication systems in the Far Eastern Branch of the Russian Academy of Sciences]. *Informatsionnye Tekhnologii i Vychislitel'nye Sistemy* [Information Technology and Computer Systems] 4:45–57.
14. Khanchuk, A. I., A. V. Konovalov, A. A. Sorokin, S. P. Korolyov, A. V. Gavrilov, V. A. Bormotov, and M. A. Serov. 2011. Instrumental'noe i informatsionno-tehnologicheskoe obespechenie seismologicheskikh nablyudenii na Dal'nem Vostoke Rossii [Instrumentation and information technology for seismological research in the Far East of Russia]. *Vestnik DVO RAN* [Bulletin of the Far Eastern Branch of the Russian Academy of Sciences] 3:127–137.
15. Sorokin, A. A., S.P. Koroljov, S.I. Smagin, and A.N. Poljakov. 2013. Maket otka-zoustoychivoy informatsionnoy sistemy dlya oblachnogo khraneniya naborov nauchnykh dannykh [Layout of a resilient information system for cloud storage of scientific data sets]. *Vychislitel'nye Tekhnologii* [Computational Technologies] 1:87–95.

Received December 2, 2013

## Contributors

**Sorokin Alexei A.** (b. 1980) — Candidate of Sciences (PhD) in technology, Head of Department, Computer Center of Far Eastern Branch of Russian Academy of Sciences, 65 Kim U Chena Str., Khabarovsk 680000, Russian Federation; alsor@febras.net

**Tarasov Alexei G.** (b. 1982) — Candidate of Sciences (PhD) in technology, senior scientist, Computer Center of the Far Eastern Branch of the Russian Academy of Sciences, 65 Kim U Chena Str., Khabarovsk 680000, Russian Federation; taleks@as.khb.ru

**Korolyov Sergei P.** (b. 1985) — scientist, Department of Information-Telecommunication Systems, Computer Center of the Far Eastern Branch of the Russian Academy of Sciences, 65 Kim U Chena Str., Khabarovsk 680000, Russian Federation; serejk@febras.net

## СИСТЕМАТИЗАЦИЯ ЗНАНИЙ О ПРОГРАММИРУЕМЫХ ЗАДАЧАХ

*A. В. Ильин<sup>1</sup>, В. Д. Ильин<sup>2</sup>*

**Аннотация:** Рассмотрены основы методологии построения символьных моделей программируемых задач, рассчитанных на применение в системах знаний о задачах. Методология основана на результатах, полученных при разработке и исследованиях систем автоматизированного конструирования программ. Для формализованного представления задач используется TSM-комплекс (Task S-Modeling) описания символьных моделей произвольных объектов в человеко-машинной среде (s-среде). Задачи представлены как конструктивные объекты (s-задачи) с определенными типами межзадачных связей и правилами конструирования. Описание s-задачи включает постановку, системы обязательных и ориентирующих правил решения, алгоритмы, библиотеки программ и тестовых примеров. В условиях применения программ указаны операционные системы, требования к аппаратным средствам и др.

**Ключевые слова:** s-моделирование; задачный конструктивный объект; s-задача; s-задачный граф; исчисление s-задач; система знаний о задачах; TSM-комплекс

**DOI:** 10.14357/08696527140314

### 1 Введение

Представление связей между понятиями в виде разрешимых задач — необходимое условие построения количественных *s-моделей* систем понятий [1] любой предметной области. В настоящее время существует огромное число разрешимых задач, запрограммированных на различных языках и реализованных в составе системных и прикладных комплексов. Какие это задачи, к каким классам относятся? Где можно найти сведения об их формулировках, методах, алгоритмах и тестовых примерах? Как использовать этот арсенал при разработке программ?

Автоматизация программирования предполагает создание удобных средств разработки и отладки программ и возможности сохранения, накопления и повторного применения ранее полученных результатов. С первых шагов автоматизации программирования вместе с языками программирования и трансляторами стали создавать различные библиотеки программ. На начальном этапе одной

---

<sup>1</sup>Институт проблем информатики Российской академии наук, avilyin@ipiran.ru

<sup>2</sup>Институт проблем информатики Российской академии наук, vdilyin@ipiran.ru

из наиболее популярных была библиотека научных подпрограмм на языке фортран (Scientific Subroutine Package — SSP) [2], разработанная IBM во второй половине 1960-х гг. В те годы SSP сыграла значительную роль в увеличении продуктивности прикладного программирования на фортране и стала своеобразным образцом. В частности, с 1970-х гг. в НИВЦ МГУ развивается библиотека программ численного анализа [3], концептуально близкая SSP. С началом реализации концепции структурного программирования [4, 5] значимость библиотек программ постоянно растет. В наши дни конкурентоспособность любой системы программирования во многом определяется ее библиотеками и удобством их применения.

Что требуется для объединения библиотек и тех программ, которые успешно работают, но не входят ни в какие библиотеки? Какой должна быть система знаний о программируемых задачах, чтобы служить основанием для конструирования программных систем? Один из возможных ответов на этот вопрос был предложен в [6, 7] применительно к порождению программ. В [8, 9] этот подход был применен в методологии конструирования параллельных программ.

Учитывая, что к настоящему времени накоплено и продолжает быстро увеличиваться число задач, алгоритмы решения которых представлены программами на различных языках и апробированы в составе системных и прикладных комплексов, целесообразно создать систему их специфицированного описания, пользуясь которой можно было бы найти сведения о постановках и алгоритмах, программах и тестовых примерах, чтобы использовать арсенал, накопленный при разработке программных средств и сервисов s-среды [1, 10]. Системы программирования целесообразно связать с системами знаний о программируемых задачах (где задачи представлены как конструктивные объекты, содержащие описание постановки, методов, алгоритмов, библиотек программ и тестовых примеров), а системы знаний — между собой [1, 6, 7, 10–12]. Структурированность, возможности расширения и коррекции таких систем станут эффективными помощниками в борьбе с пресловутой проблемой сложности.

Цель этой статьи — представить основы методологии формализованного описания программируемых задач, рассчитанной на разработку систем знаний о задачах.

## 2 О средствах описания моделей программируемых задач

Для формализованного представления s-моделей программируемых задач использован *TSM-комплекс средств описания s-моделей* [12]. Первые версии комплекса были разработаны в конце 1980-х гг. [6, 7]. В [10–12] комплекс был усовершенствован. Далее представлена часть средств TSM-комплекса, используемая в этой статье.

Для выделения определений, замечаний, примеров, имен понятий и отдельных частей описания используются следующие средства:

$\square$  (фрагмент описания)  $\square \approx$  утверждение (определение, аксиома и др.) (здесь и далее символ  $\approx$  заменяет слово «означает»);

$\diamondsuit$  (фрагмент описания)  $\diamondsuit \approx$  замечание;

$\circlearrowleft$  (фрагмент описания)  $\circlearrowleft \approx$  пример.

Курсивом выделены первые вхождения названий понятий [определеняемых или определенных (последние могут быть гиперссылками)] и фрагменты описания, к которым авторы хотят привлечь внимание.

Для теоретико-множественных и других формул применяется одноуровневая форма записи. Не накладывается ограничений на число индексов и помечающих символов (помет) для переменных. Все индексы и пометы записываются в строчку внутри квадратных скобок, следующих сразу за индексируемой (или / и помеченной) переменной. Индексы, определяющие элемент массива, отделяются запятыми, индексированные индексы — косой чертой «/». Верхний индекс от нижнего отделяется точкой с запятой. Если в описании индекса точка с запятой не встречается, то индекс считается нижним. Если сразу после точки с запятой стоит закрывающая квадратная скобка, то задан только верхний индекс.

$\circlearrowleft x[\text{out}; j = 1 \dots n]$   $\approx$  вектор  $x$  из  $n$  компонент, имеющий помету  $\text{out}$ ;

$a[\text{inp}; i = 1 \dots m, j = 1 \dots n]$   $\approx$  матрица  $a$  размера  $m \times n$ , имеющая помету  $\text{inp}$ ;

$c['; 1]$   $\approx$   $c$ -один со штрихом (штрих «'» — верхняя помета; 1 — нижний индекс);

$d[j;]$   $\approx$   $d$  с верхним индексом  $j$ ;

$d[i/j]$   $\approx$   $d$  с нижним индексом « $i$   $j$ -е»  $\circlearrowleft$ .

Аргументы функции размещаются в круглых скобках, стоящих сразу за идентификатором, обозначающим функцию.

$\circlearrowleft f(x)$   $\approx$   $f$  от  $x$ ;  $f[\max;](x[i = 1 \dots n])$   $\approx$   $f$  с верхней пометой  $\max$  от  $x[i = 1 \dots n]$ .  $\circlearrowleft$

Теоретико-множественные:

$a : \text{elem } A \approx a$  является элементом множества  $A$ ;

$A < B \approx A$  — подмножество  $B$  (когда оговорено, что  $A$  и  $B$  рассматриваются как множества);

$C \leq B \approx C$  является подмножеством  $B$  или совпадает с ним;

$A \vee B \approx$  объединение множеств  $A$  и  $B$ ;

$A \wedge B \approx$  пересечение множеств  $A$  и  $B$ ;

$A \setminus B \approx$  разность множеств  $A$  и  $B$ ;

$A * B \approx$  декартово произведение множеств  $A$  и  $B$ ;

$R \leq A * B \approx$  бинарное отношение, заданное на множествах  $A$  и  $B$ .

$\square T\text{up } X \approx$  множество  $X$ , элементы которого имеют фиксированный набор атрибутов и семейство допустимых операций. Может иметь подтипы, называемые специализациями типа  $X$ , и надтипы, называемые обобщениями типа  $X$ .  $\square$

$\square$  Специализация типа  $X$  — порождение подтипа  $X :: \text{rule}$  (здесь сдвоенное двоеточие — символ специализации) с семейством связей, расширенным добав-

лением правила rule. Специализацией называем и результат этого порождения  $X :: rule$ .

$X :: (rule1) :: rule2$  — специализация типа  $X :: rule1$  по связи rule2. Число специализирующих связей в последовательности не ограничено. При этом правила, предшествующие последнему, заключены в круглые скобки, а перед открывающей скобкой каждой пары скобок — сдвоенное двоеточие.  $\square$

$\square$  *Обобщение типа Z* — это порождение его надтипа  $Z \# rule$  или  $Z \# (rule \leftarrow rule G)$  путем исключения правила rule из семейства связей, соответствующего типу  $Z$ , или замены rule на более общее правило rule  $G$  (здесь «решетка» — символ обобщения).  $\square$

Комплекс TSM рассчитан на создание строчных одноуровневых описаний s-моделей без использования специализированных редакторов.

### 3 Символьное моделирование задач

$\square$  *S-задача* — это четверка  $\{Formul, Rulsys, Alg, Prog\}$ , где Formul — постановка задачи; Rulsys — множество систем *обязательных* и *ориентирующих* правил решения задачи [13], поставленных в соответствие Formul; Alg — объединение множеств алгоритмов, каждое из которых соответствует одному элементу из Rulsys; Prog — объединение множеств программ, каждое из которых поставлено в соответствие одному из элементов Alg. Постановка задачи Formul — это пара  $\{\text{Mem}, \text{Rel}\}$ , где Mem — множество понятий задачи, на котором задано разбиение  $\text{Mem} = \text{Inp} \vee \text{Out}$  ( $\text{Inp} \wedge \text{Out} = 0$ ) и совокупность Rel связей между понятиями, определяющая бинарное отношение  $\text{Rel} < \text{Inp} * \text{Out}$ . Множество Mem назовем также *памятью задачи*, а Inp и Out — ее *входом* и *выходом*, значения которых предполагается соответственно задавать и искать.  $\square$

Для каждого элемента из Rulsys, Alg и Prog задано описание применения. Описания применения элементов Rulsys включают спецификацию типа решателя задачи (автономная s-машина [1], сетевая коопeração s-машин, коопे-рация человек–s-машина и т. п.), требования к информационной безопасности и др. Описания применения элементов из Alg включают данные о допустимых режимах работы решателя задачи (автоматический локальный, автоматический распределенный, интерактивный локальный и др.), о требованиях к полученному результату и др. Описания применения программ включают данные о языках реализации, операционных системах и др.

$\diamond$  Каждая программа сопровождается ссылками на наборы тестовых примеров [10–12].  $\diamond$

В общем случае множества Rulsys, Alg и Prog могут быть пустыми: число их элементов зависит от степени изученности задачи.

$\square$  *S-алгоритм* — система правил решения задачи (соответствующая одному из элементов Rulsys), позволяющая за конечное число шагов поставить в однозначное соответствие заданному набору данных, принадлежащему Inp, результирующий набор, принадлежащий Out.

Выполнение s-алгоритма подразумевает следующие действия:

- распознавание набора входных данных (определение его принадлежности множеству  $\text{Inp}$ ): если набор принадлежит  $\text{Inp}$ , то переход к следующему пункту, в противном случае — остановка;
- интерпретацию набора из  $\text{Inp}$  (получение результирующего набора данных, принадлежащего  $\text{Out}$ );
- запись результирующего набора данных в заданную область памяти;
- остановку.  $\square$

$\square$  *S-программа* — реализованный (на языке программирования высокого уровня, машинно-ориентированном языке и/или в системе машинных команд) s-алгоритм, представленный в форме сообщения, определяющего поведение s-машинного решателя задачи с заданными свойствами. Существует в символичном, кодовом и сигнальном воплощениях, связанных отношениями трансляции.  $\square$

$\square$  *Спецификация spec s-задачи* — это пара ( $\text{Formul}$ ,  $\text{as}$ ), где  $\text{as}$  — описание применения.  $\square$

$\bigcirc$  TSM-запись формулировки задачи линейного программирования:  
вход задачи  $\text{Inp} = \{\text{матрица } a[i = 1 \dots m, j = 1 \dots n] \text{ коэффициентов ограничений, вектор } b[i = 1 \dots m] \text{ правых частей ограничений, вектор } c[j = 1 \dots n] \text{ коэффициентов целевой функции};\}$   
выход  $\text{Out} = \{\text{искомый вектор } x[\max; j = 1 \dots n]\}.$

Правило  $\text{rul}$  максимизации по  $x[j = 1 \dots n]$  целевой функции  $c[j = 1 \dots n] * x[j = 1 \dots n]$  при ограничениях  $a[i = 1 \dots m, j = 1 \dots n] * x[j = 1 \dots n] \leq b[i = 1 \dots m]$  и  $x[j = 1 \dots n] \geq 0$  имеет следующий вид:

$$\max[x[j = 1 \dots n] :: (a[i = 1 \dots m, j = 1 \dots n] * x[j = 1 \dots n] \leq b[i = 1 \dots m]) :: x[j = 1 \dots n] \geq 0] (c[j = 1 \dots n] * x[j = 1 \dots n]). \bigcirc$$

### 3.1 Связи по памяти

Связи по памяти между s-задачами определяются тремя типами функций, каждая из которых является функцией двух аргументов и позволяет поставить в соответствие паре s-задач некоторую третью s-задачу, образованную из этой пары.

$\square$  S-задача  $a$  связана с s-задачей  $b$  по памяти, если существует хотя бы одна пара элементов  $\{\text{elem } \text{Mem}[a], \text{elem } \text{Mem}[b]\}$ , принадлежащих памяти  $\text{Mem}[a]$  s-задачи  $a$  и памяти  $\text{Mem}[b]$  s-задачи  $b$ , относительно которой определено общее означивание (элементы имеют одно и то же множество значений). Пусть  $S$  и  $H$  — множества s-задач и  $D \leq S * S$ . Если каждой паре  $\{s[i], s[j]\}$  элементов из  $D$  ставится в соответствие определенный элемент из  $H$ , то будем говорить, что задана *функция связи по памяти*  $h = \text{conn}(s[i], s[j])$ . При этом  $D$  будем

называть областью определения функции  $\text{conn}$  и обозначать  $D[\text{conn}]$ . Множество  $R = \{h : \text{elem } H; h = \text{conn}(s[i], s[j]); s[i] : \text{elem } D[\text{conn}], s[j] : \text{elem } D[\text{conn}]\}$  называется областью значений функции  $\text{conn}$ .  $\square$

Тип связи зависит от содержимого пересечения по памяти: составлена ли связь из элементов выхода одной и входа другой задачи, из элементов выходов задач или из элементов их входов или же связь получена путем комбинации предыдущих способов.

Будем обозначать функцию связи по памяти типа *вход–вход* через  $\text{conn}[x]$ , *выход–вход* — через  $\text{conn}[yx]$  и *выход–выход* — через  $\text{conn}[y]$ .

### 3.2 Специализация и обобщение

*S-задача* может быть прообразом некоторого непустого множества *s-задач* или образом некоторого прообраза либо быть одновременно и образом какой-то одной *s-задачи*, и прообразом некоторого множества других *s-задач*.

Предусмотрены следующие типы родовых связей между *s-задачами*:

- *s-(специализация задачи)* — указание на *s-задачу*  $b$ , частную по отношению к исходной *s-задаче*  $a$ :  $b = a::\text{rule}$ ;
- *s-(обобщение задачи)* — указание на *s-задачу*  $b$ , которая служит обобщением исходной *s-задачи*  $a$ :  $b = a#\text{rule}$  или  $b = a\#(\text{rule} \leftarrow \text{rule } G)$ .

### 3.3 Конструирование

$\square$  *S-(конструирование задачи)* реализуется посредством связи по памяти между задачами. Элементарная задачная конструкция — это задачная пара. Из задачных пар можно построить более сложную конструкцию, если рассматривать их как задачные элементы. Любая задачная конструкция, в свою очередь, может быть использована как составляющая еще более сложной задачной конструкции.  $\square$

$\square$  *S-задача* называется *атомарной*, если ее формулировка не представлена в виде структуры, заданной на некотором множестве формулировок других *s-задач*.  $\square$

Будем также говорить об атомарной *s-задаче* как о *простой s-задаче*. Простая задача (с точки зрения построителя задачных конструкций) не наделена внутренней структурой и потому не подлежит декомпозиции.

Каждая задачная конструкция может быть объявлена некоторой новой задачей. В свою очередь, эти новые задачи вместе с атомарными могут быть применены при конструировании задач.

### 3.4 Конкретизация

$\square$  *S-(конкретизация задачи)* — это переход (формулировка  $\rightarrow$  система обязательных и ориентирующих правил решения задачи  $\rightarrow$  алгоритм  $\rightarrow$  заготовка программы  $\rightarrow$  программа).  $\square$

Для s-задач, имеющих пустое множество программ ( $\text{Prog} = 0$ ), конкретизация завершается выбором или разработкой алгоритма. Если и  $\text{Alg} = 0$ , s-задача может быть использована в s-(конструировании задач), но не может быть конкретизирована.

## 4 Система знаний об s-задачах

□ *Система pS знаний о задачных конструктивных объектах* (называемых также *p-объектами*) — это триада  $\langle \text{pA}, \text{lng}, \text{intr} \rangle$ , где  $\text{pA}$  — задачная область,  $\text{lng}$  — язык спецификации *p-объектов*,  $\text{intr}$  — интерпретатор спецификаций искомых *p-объектов* на  $\text{pA}$ . □

### 4.1 Модель задачной области

Пусть  $P$  — множество всех *p-объектов*, а  $A < P$  — его непустое подмножество. При этом в  $A$  (содержащем не менее двух элементов) не существует ни одного элемента, который не был бы связан по памяти хотя бы с одним элементом из  $A$ .

□ *S-модель*  $\text{ra}$  задачной области  $\text{pA}$  — это *p-объект*, который задается парой  $\langle \text{память tem}[A] \text{ множества задач } A \text{ задачной области } \text{pA}, \langle \text{семейство rul}(tem[A]) \text{ связей, заданных на tem}[A] \rangle \rangle$ .

Непустое множество  $\text{tem}[A]$  элементов памяти разбито на три подмножества: входов  $\text{inp}[A]$  задач, выходов  $\text{out}[A]$  задач и подмножество  $\text{og}[A]$ , каждый из элементов которого является и входом, и выходом некоторых задач. Любое одно из этих подмножеств может быть пустым; могут быть одновременно пустыми  $\text{inp}[A]$  и  $\text{out}[A]$ . □

В отличие от памяти задачи, состоящей из входа и выхода, память задачной области содержит подмножество  $\text{og}$  элементов памяти, каждый из которых может быть или задан (как входной), или вычислен (как выходной). Будем называть такие элементы памяти *обратимыми*, а  $\text{og}$  — подмножеством обратимых элементов. Подмножество  $\text{inp}$  будем называть подмножеством *задаваемых*, а подмножество  $\text{out}$  — подмножеством *вычисляемых* элементов.

Задачная область  $\text{pA}$  служит *s-моделью* [1, 12], на которой интерпретируются спецификации искомых задач, составленные на языке  $\text{lng}$ .

### 4.2 Интерпретация и разрешающая структура

◊ Интерпретация заключается в постановке в соответствие некоторому подмножеству (или паре подмножеств) памяти  $\text{tem}[A]$  некоторой подобласти задачной области  $\text{pA}$ , названной *разрешающей структурой*. Интерпретация спецификации искомого *p-объекта* на  $\text{pA}$  — это конструктивное доказательство существования разрешающей структуры. ◊

### 4.3 Задачные графы

□ *Задачный граф* служит представлением задачной области, рассчитанным на реализацию процесса р-конструирования и формализацию задачных знаний. Множество вершин графа составлено из задачных объектов. Оно называется *задачным базисом графа* и обозначается  $p\text{-basis}$ . Ребро задачного графа — это пара вершин с непустым пересечением по памяти. Нагрузка ребра определяется множеством всех пар элементов памяти, входящих в это пересечение. Каждая вершина графа имеет память. Память вершины — это память задачи (или задачной области), которую представляет вершина. □

□ *Составная задача* comp — подобласть задачной области  $pA$ , которая содержит не менее двух элементов из множества задач  $A$  и на памяти которой задано разбиение:  $\text{mem}[\text{comp}] = \text{inp}[\text{comp}] \vee \text{out}[\text{comp}]$ ;  $\text{inp}[\text{comp}] \wedge \text{out}[\text{comp}] = 0$ , определяющее вход  $\text{inp}[\text{comp}]$  и выход  $\text{out}[\text{comp}]$  составной задачи. Составной задаче поставлен в соответствие ориентированный граф, вершинами которого являются задачи. Каждая вершина помечена именем задачи. Ребра графа — это пары задач с непустыми пересечениями по памяти. □

◊ Составная задача может быть построена путем последовательного применения функций связи по памяти. ◊

□ В зависимости от состава вершин определим следующие типы задачных графов:

- *U-граф* имеет множество вершин только из простых задач;
- в *C-графе* хотя бы одна вершина представлена составной задачей и нет вершин, представляющих собой задачную область;
- в *G-графе* не менее одной вершины представлено задачной областью (остальные могут быть простыми и составными задачами).

Связный граф с непустым множеством ребер и задачным базисом  $p\text{-basis}$ , все элементы которого являются простыми задачами, называется *U-графом* и обозначается *U-graph*:  $U\text{-graph} = (p\text{-basis}, \text{set}[ver])$ , где  $\text{set}[ver]$  — множество ребер задачного графа. Объединение памяти задачных вершин, составляющих базис, называется памятью *U-графа* и обозначается  $\text{mem}[U\text{-graph}]$ .

На памяти *U-графа* задано разбиение на подмножества:

- $\text{Giv}[U\text{-graph}]$  задаваемых (входных) элементов памяти;
- $\text{Comput}[U\text{-graph}]$  вычисляемых (выходных) элементов памяти;
- $\text{Or}[U\text{-graph}]$  обратимых элементов памяти (определяется разностью  $\text{mem}[U\text{-graph}] \setminus (\text{Giv}[U\text{-graph}] \vee \text{Comput}[U\text{-graph}])$ ).

Память *C-графа* и память *G-графа* обозначаются и определяются аналогично памяти *U-графа*. □

#### **4.4 G-графы — средство формализации знаний о р-объектах**

Система знаний об s-задачах обеспечивает процессы s-(*специализации, конкретизации и конструирования*).

◊ Возможность существования в задачном графе одного или нескольких узлов, являющихся задачными областями, имеет принципиальное значение для формализации задачных знаний. ◊

Конкретным воплощением задачной области может быть граф любого типа (U-, C- или G-граф). Тот факт, что G-граф может замещать задачный узел, предоставляет возможность произвольного усложнения задачной области. Она может быть представлена, в частности, посредством G-графа, базис которого состоит только из вершин, представленных G-графами. Построение множества задачных конструктивных объектов в составе системы знаний об s-задачах основано на *исчислении р-объектов* [6–12, 14, 15].

### **5 Прикладное значение**

Конструирование разрешающих структур на задачных графах [7, 11, 12] позволяет путем интерпретации интерактивно формируемой спецификации задачи получить соответствующую ей структуру разрешимых задач. Реализация системы знаний об s-задачах и методологии конструирования разрешающих структур значительно повышает эффективность разработки программного обеспечения.

◊ Применительно к конструированию разрешающих структур на задачных графах существование набора тестовых примеров для каждого р-объекта — необходимое условие реализации методологии. Заметим, что тестирование полученной разрешающей структуры входит в конструктивное доказательство ее существования. ◊

Нынешние библиотеки программ различного назначения не всегда удовлетворяют требованиям, предъявляемым к s-модели задачного конструктивного объекта. Часто они не содержат ссылок на наборы тестовых примеров. Нет возможности протестировать составляющие нынешних операционных систем и работающих под их управлением приложений. Правило продавать программные продукты с заранее оговоренной возможностью доступа к наборам тестовых примеров (выложенных на сайте производителя) заметно уменьшит число ошибок и уязвимостей на этапе разработки и увеличит вероятность их выявления после начала эксплуатации. Известно, что составление наборов тестовых примеров требует понимания не только сути задачи, выполняемой тестируемой программой, но и допустимых условий ее применения. Если правило предъявлять потребителю наборы тестовых примеров станет обязательным, изменится отношение разработчиков и к изготовлению таких наборов, и к производству программных продуктов.

## 6 Заключение

Рассмотренные основы методологии систематизации символьных моделей программируемых задач рассчитаны на построение систем знаний о задачах. Методология применима для представления знаний о задачах не только в системах автоматизированного конструирования программ [7, 11, 12], при разработке и исследованиях которых она сформировалась, но и для представления знаний о задачах в системах автоматизированного проектирования любого назначения. Накоплен опыт успешного применения этой методологии при разработке приложений [13] и формировании образовательных ресурсов [10].

## Литература

1. Ильин А.В., Ильин В.Д. Основы теории s-моделирования. — М.: ИПИ РАН, 2009. 143 с. <http://smodeling.files.wordpress.com/2010/01/basics-theory-sm-20092.pdf>.
2. Sweger K. 1130 Scientific Subroutine Package (1130-CM-02X): Programmer's manual. — IBM, 1967. <http://media.ibm1130.org/1130-006-ocr.pdf>.
3. Библиотека численного анализа НИВЦ МГУ. [http://num-anal.srcc.msu.su/lib\\_na/libnal.htm](http://num-anal.srcc.msu.su/lib_na/libnal.htm).
4. Dijkstra E. Structured programming // Software engineering techniques / Eds. J. Burton, B. Randell. — Brussels: NATO Science Committee, 1969. P. 89–93.
5. Wirth N. Program development by stepwise refinement // Comm. ACM, 1971. Vol. 14. No. 4. P. 221–227.
6. Ильин В. Д. Представление знаний о задачах в системе порождения программ. — М.: ИПИ АН, 1989. 50 с.
7. Ильин В. Д. Система порождения программ. — М.: Наука, 1989. 264 с.
8. Ilyin V. D. PPS-construction: A methodology for computer aided construction of parallel program systems // Software for Multiprocessors and Supercomputers: Theory, practice, experience: 2nd Conference (International) Proceedings. — Moscow: Institute for System Programming of the Russian Academy of Sciences, 1994. P. 356–364.
9. Ilyin V. D. A methodology for knowledge based engineering of parallel program systems // Industrial and Engineering Applications of Artificial Intelligence and Expert Systems: 8th Conference (International) Proceedings. — Melbourne: Gordon and Breach Publs., 1995. P. 805–809.
10. Ильин А. В., Ильин В. Д. Научно-образовательные веб-ресурсы. S-моделирование. — М.: ИПИ РАН, 2013. 112 с. <http://smodeling.files.wordpress.com/2013/12/nowrsm-2013.pdf>.
11. Ильин А. В. Конструирование разрешающих структур на задачных графах системы знаний о программируемых задачах // Информационные технологии и вычислительные системы, 2007. № 3. С. 30–36.
12. Ильин В. Д. Система порождения программ. Версия 2013 г. — М.: ИПИ РАН, 2013. 142 с. <http://smodeling.files.wordpress.com/2013/12/vd-ilyin-spp-2013.pdf>.
13. Ilyin A. V., Ilyin V. D. The technology of interactive resource allocation in accordance with the customizable system of rules // Appl. Math. Sci., 2013. Vol. 7. No. 143. P. 7105–7111. doi: 10.12988/ams.2013.311649.

14. Gavrilenko Yu. V., Ilyin V. D. Spaces of local objects for semantic interpretation of problems. I // J. Comput. Sys. Sc. Int., 1996. Vol. 35. No. 5. P. 687–694.
15. Gavrilenko Yu. V., Ilyin V. D. Calculus of task-constructive objects and their interpretations. II // J. Comput. Sys. Sc. Int., 1997. Vol. 36. No. 5. P. 704–712.

Поступила в редакцию 16.07.14

---

## SYSTEMATIZATION OF KNOWLEDGE ABOUT PROGRAMMABLE TASKS

A. V. Ilyin and V. D. Ilyin

Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str.,  
Moscow 119333, Russian Federation

**Abstract:** The article describes basics of the methodology for systematization of symbolic models of programmable tasks intended for development of systems of knowledge about tasks. This methodology is based on the results achieved in development and research of systems of computer-aided software design. The TSM-complex (Task S-Modeling) for description of models of arbitrary objects in human-machine environment (s-environment) is used for formalized representation of tasks. The tasks are represented as constructive objects (s-tasks) having certain types of relations and rules of design. S-task specification includes formulation, systems of mandatory and adjusting rules for solving, algorithms, sets of programs, and test cases. Descriptions of program applications include information about operating systems, hardware requirements, etc.

**Keywords:** s-modeling; task constructive object; s-task graph; system of knowledge about tasks

**DOI:** 10.14357/08696527140314

## References

1. Ilyin, A. V., and V. D. Ilyin. 2009. *Osnovy teorii s-modelirovaniya* [Basics of the theory of s-modeling]. Moscow: Institute of Informatics Problems of RAS. 143 p. Available at: <http://smodeling.files.wordpress.com/2010/01/basics-theory-sm-20092.pdf> (accessed March 27, 2014).
2. IBM. 1967. *1130 Scientific Subroutine Package (1130-CM-02X) Programmer's Manual*. Available at: <http://media.ibm1130.org/1130-006-ocr.pdf> (accessed March 27, 2014).
3. Biblioteka chislennogo analiza NIVTs MGU [Library of numerical analysis of the Scientific Research Computing Center of the Lomonosov Moscow State University]. Available at: [http://num-anal.srcc.msu.su/lib\\_na/libnal.htm](http://num-anal.srcc.msu.su/lib_na/libnal.htm) (accessed March 27, 2014).

4. Dijkstra, E. 1969. Structured programming. *Software engineering techniques*. Eds. J. Burton and B. Randell. Brussels: NATO Science Committee. 89–93.
5. Wirth, N. 1971. Program development by stepwise refinement. *Comm. ACM.* 14(4):221–227.
6. Ilyin, V. D. 1989. *Predstavlenie znaniy o zadachakh v sisteme porozhdeniya programm* [Representation of knowledge about tasks in the system of program generating]. Moscow: Institute of Informatics Problems of the Academy of Sciences of the USSR. 50 p.
7. Ilyin, V. D. 1989. *Sistema porozhdeniya programm* [The system of program generating]. Moscow: Nauka. 264 p.
8. Ilyin, V. D. 1994. PPS-construction: A methodology for computer aided construction of parallel program systems. *2nd Conference (International) "Software for Multiprocessors and Supercomputers. Theory, practice, experience" Proceedings.* 356–364.
9. Ilyin, V. D. 1995. A methodology for knowledge based engineering of parallel program systems. *8th Conference (International) "Industrial and Engineering Applications of Artificial Intelligence and Expert Systems" Proceedings.* 805–809.
10. Ilyin, A. V., and V. D. Ilyin. 2013. *Nauchno-obrazovatel'nye veb-resursy. S-modelirovaniye* [Scientific-educational Web-resources. S-modeling]. Moscow: Institute of Informatics Problems of RAS. 112 p. Available at: <http://smodeling.files.wordpress.com/2013/12/nowrsm-2013.pdf> (accessed March 27, 2014).
11. Ilyin, A. V. 2007. Konstruirovaniye razreshayushchikh struktur na zadachnykh grafiakh sistemy znaniy o programmiruemym zadachakh [Construction of solving structures in task graphs of system of knowledge about programmable tasks]. *Informatsionnye Tekhnologii i Vychislitel'nye Sistemy* [Information Technology and Computing] 3:30–36.
12. Ilyin, V. D. 2013. *Sistema porozhdeniya programm. Versiya 2013 g.* [The system of program generating. Version 2013]. Moscow: Institute of Informatics Problems of RAS. 142 p. Available at: <http://smodeling.files.wordpress.com/2013/12/vd-ilyin-spp-2013.pdf> (accessed March 27, 2014).
13. Ilyin, A. V., and V. D. Ilyin. 2013. The technology of interactive resource allocation in accordance with the customizable system of rules. *Appl. Math. Sci.* 7(143):7105–7111. doi: 10.12988/ams.2013.311649.
14. Gavrilenko, Yu. V., and V. D. Ilyin. 1996. Spaces of local objects for semantic interpretation of problems. I. *J. Comput. Sys. Sc. Int.* 35(5):687–694.
15. Gavrilenko, Yu. V., and V. D. Ilyin. Calculus of task-constructive objects and their interpretations. II. *J. Comput. Sys. Sc. Int.* 36(5):704–712.

*Received July 16, 2014*

## Contributors

**Ilyin Alexander V.** (b. 1975)—Candidate of science (PhD) in technology, senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; avilyin@ipiran.ru

**Ilyin Vladimir D.** (b. 1937)—Doctor of Science in technology, Head of Laboratory, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vdilyin@ipiran.ru

## ИСПОЛЬЗОВАНИЕ ХЕШ-ФУНКЦИЙ ДЛЯ ПОВЫШЕНИЯ СКОРОСТИ МОРФОЛОГИЧЕСКОГО АНАЛИЗА РУССКИХ ТЕКСТОВ

*Н. В. Сомин<sup>1</sup>, М. М. Шарнин<sup>2</sup>*

**Аннотация:** Рассматривается проблема повышения эффективности морфологического анализа русского языка. Описывается разработанная в ИПИ РАН система морфологического анализа: набор морфологических характеристик и алгоритмы работы. Указываются задачи и системы, связанные с проблемой логико-семантического анализа текстов, в которых данная морфологическая система нашла применение. Обсуждаются особенности системы с точки зрения занимаемой памяти и скорости работы. Предлагается способ хранения морфо-лексической информации с помощью хеш-функций, обеспечивающих высокую скорость доступа. Обсуждаются трудности, возникающие при реализации такого подхода, и рассматриваются пути их преодоления. Приводится структура информационных массивов новой версии и реализованные в ней поисковые алгоритмы, а также даются сведения по подсистеме ввода и корректировки морфологической информации. Приводятся конкретные параметры новой реализации морфологической системы и данные по ускорению работы по сравнению с предыдущей версией. В заключение обсуждаются возможности по развитию новой версии морфологии и перенесению предложенного подхода к реализации на другие компоненты лингвистического процессора.

**Ключевые слова:** морфологический анализ; лингвистический процессор; морфологическая омонимия; хеш-функция

**DOI:** 10.14357/08696527140315

### 1 Введение

Задача морфологического анализатора — нормализация слов, наиболее полное и точное определение морфологических признаков лексем, а также (в ряде случаев) нахождение их семантического класса. Отметим, что к настоящему времени разработан целый ряд морфологических анализаторов русского языка, среди которых можно отметить [1–3]. Авторы принимали участие в создании разработанной в ИПИ РАН системы морфологического анализа [4], а также ее интеграции в состав целого ряда лингвистических процессоров:

---

<sup>1</sup>Институт проблем информатики Российской академии наук, somin@post.ru

<sup>2</sup>Институт проблем информатики Российской академии наук, mc@keywen.com

«Криминал», «Аналитик», «Поток», «Лингвопроцессор», «Лингво-ИИ» [5–11] и ряда других. Помимо этого система морфологического анализа применялась при реализации экспериментальных лингвистических процессоров ЭССЕИСТ и ТЕРМИН-3 [12], а также в системах рубрицирования текстов [13, 14]. Опыт эксплуатации морфологической системы отражен в работе [4]. Везде входной поток текстов, проходя через морфологический блок (интегрированный с блоком лексического анализа), дополнялся подробной морфо-лексической информацией, на основе которой уже производился дальнейший анализ.

Однако в связи с доступностью на современном этапе огромного количества текстов из Интернета и, соответственно, необходимостью обработки больших объемов данных (Big Data) к системе морфологического анализа предъявляются новые требования. В частности — требование высокой производительности. Причем, как представляется, уповать на простое увеличение скорости работы современных процессоров нет оснований, поскольку темпы роста их вычислительной мощности явно отстают от темпов увеличения объемов текстов, требующих обработки. Поэтому необходимо искать другие пути повышения эффективности морфологического анализа. Они были найдены в использовании ресурса оперативной памяти современных компьютеров, который за последнее десятилетие вырос более чем на порядок. Именно за счет использования большего объема памяти в новой версии морфологической системы удалось существенно повысить скорость анализа, не потеряв (и даже увеличив) функциональные возможности старой версии системы.

## 2 Возможности системы морфологического анализа

Разработанная система морфологического анализа позволяет обрабатывать текст на русском и английском языках, выполняя для каждого слова определение его канонической формы и выдавая его полный морфологический анализ. Кроме того, система позволяет:

- выполнять генерирование всей парадигмы словоформ по заданной словоформе;
- формировать предметные терминологические словари и идентифицировать термины с учетом парадигматических изменений [15].

Ниже представлен набор морфологических признаков с указанием соответствующих символов (на первом месте) и формируемых на их основе фрагментов-признаков (в скобках — они используются программами семантического анализа на языке ДЕКЛ [6]):

f	—	фамилия (FAM);
i	—	имя (NAME_);
h	—	отчество (NAME_2);
N	—	существительное (NOUN);
A	—	прилагательное (ADJ);

.	единственное число (по умолчанию);
:	множественное число (MANY);
м	мужской род;
ж	женский род;
с	средний род;
и	именительный;
р	родительный (КОГО);
д	дательный (КОМУ);
в	винительный;
т	творительный (КЕМ);
п	предложный;
D	деепричастие (DPRICH);
R	наречие (ADJ);
G	междометие (MZ);
U	числительное (NUM_W);
O	отглагольное прилагательное типа «былый», «бытий» (ADJ_O);
H	частица (PART);
P	предлог (PREP);
M	местоимение (PRON);
T	причастие (PRICH);
V	глагол (VERB);
S	союз (CONJ);
E	модальное слово (VERB_2);
Z	указательное местоимение «тот», «этот» (PRON);
W	вводное слово типа «надо», . . . (AUX);
=	настоящее;
>	будущее;
<	прошедшее;
+	положительная степень (по умолчанию);
*	сравнительная степень;
\$	превосходная степень;
л	1-е лицо;
Л	2-е лицо;
L	3-е лицо;
П	полная форма (по умолчанию — FULL);
к	краткая форма (SHORT);
Г	одновременно существительное + прилагательное, числительное, причастие («малый», «нищий»);
ф	инфinitив (INFIN);
ъ	изъявительное наклонение;
!	повелительное наклонение;
я	форма без -ся/-сь (по умолчанию);
I	одушевленное;

X	—	неодушевленное;
)	—	безличный глагол;
{	—	личный/безличный глагол;
[	—	личное местоимение;
v	—	личный глагол;
}	—	личный/безличный глагол;
/	—	совершенный/несовершенный вид;
ш	—	совершенный вид;
н	—	несовершенный вид;
-	—	отрицательное местоимение;
F	—	определительное местоимение;
Ж	—	притяжательное местоимение;
q	—	порядковое числительное;
Q	—	количественное числительное;
?	—	относительное (вопросительное) местоимение;
s	—	собирательное местоимение;
ч	—	непереходный глагол;
~	—	переходный глагол;
@	—	переходный/непереходный глагол;
w	—	слово в значении сказуемого (например, « <i>цен</i> »);
u	—	неопределенное местоимение (« <i>некоторый</i> »);
v	—	возвратный глагол (с <i>ся</i> / <i>съ</i> );
d	—	действительный залог;
b	—	страдательный залог;
ю	—	разносклоняемое;
Б	—	архаичное;
Д	—	числительное, характеризуется числом;
И	—	частица без ударения;
Й	—	ударный предлог;
Н	—	1-е склонение;
Ф	—	2-е склонение;
Ц	—	3-е склонение;
З	—	группа местоимения;
Ч	—	1-е спряжение;
Ш	—	2-е спряжение;
ы	—	инициалы (NAME!);
6	—	в слове более одной прописной (HEAD_1);
7	—	слово начинается с прописной (NAME0);
0	—	в слове все прописные (HEAD);
э	—	английское слово (ENGL);
8	—	целое число (NUM);
9	—	число с точкой (NUM_F);
щ	—	сложный предлог;

t	— форма английского глагола <i>be</i> ;
R	— артикль;
Я	— личное местоимение (PRON_1);
#	— морфология определена по аналогии;
a	— англ. <i>ing</i> -форма (WEB_C);
b	— англ. <i>ed</i> -форма past-1 (MAIL_E);
E	— англ. <i>ed</i> -форма past-2.

Для одной лексемы блок морфологического анализа выдает множество признаков, набор которых и характеризует морфологический тип. Кроме морфологических блок выдает еще множество лексических признаков, о которых речь шла выше, а также ряд фонетических, которые могут быть использованы для синтеза речи.

Особое место занимает признак «#». Он означает, что данный набор признаков сформирован «по аналогии», т. е. была найдена словоформа с окончанием (т. е. несколькими конечными буквами), таким же как у данной лексемы, и морфологические признаки этой словоформы были приписаны данной лексеме. Варианты разбора «по аналогии» применяются для лексем, которых нет в морфологическом словаре.

### **3 Борьба с морфологической омонимией**

Использование обобщенного словаря основ (ОСО) может приводить к лишним вариантам морфологического анализа. Было предложено два эффективных способа борьбы с морфологической омонимией.

Первый способ — эмпирический алгоритм, отбрасывающий наименее вероятные варианты морфологического анализа. Такая «зачистка» вариантов выполняется по многим критериям, учитывающим наличие слова в словаре основ (СО), длину основы со словарем хвостов основ (СХО), часть речи. Эмпирический алгоритм фиксирует все варианты разбора в порядке их вероятности, что необходимо для ряда приложений, когда используется только один вариант морфологического анализа.

Второй способ — частичный синтаксический анализ. В предложении слово вступает в синтаксические связи с другими словами, и выявление этих связей позволяет отбросить варианты морфологического анализа, этим связям не удовлетворяющие. Прежде всего было реализовано распознавание двух конструкций: полного согласования и генетической цепочки.

Алгоритмы борьбы с морфологической омонимией являются основным содержанием так называемого процесса «постморфологии», т. е. обработки результатов морфологического анализа на завершающем этапе. Как показала практика, эти процессы занимают в морфологической системе большое место — как по алгоритмической сложности, так и по времени их выполнения.

## 4 Особенности блока английской морфологии

Английский морфологический словарь использует разработанное для русской морфологии программное обеспечение, которое оказалось возможным адаптировать к специфике английского языка. Блоки английской и русской морфологии выдают практически одни и те же морфологические характеристики, что позволяет использовать для синтаксико-семантического анализа англоязычных текстов те же средства, что и для русского языка. В результате появилась возможность записывать лингвистические знания для этих языков в одном и том же формализме.

Общий объем словаря основ блока английской морфологии — около 85 тыс. слов. Для повышения качества работы этого блока в него был введен ряд специфических для английского языка алгоритмов, которые в основном касаются отсеяния лишних вариантов морфологического анализа, поскольку слова английского языка чрезвычайно омонимичны — очень часто одно и то же слово может быть и существительным, и глаголом, и прилагательным. В блоке английской морфологии реализованы алгоритмы, позволяющие корректно отбрасывать лишние варианты (другие варианты отсеиваются в процессе синтактико-семантического анализа). Блок лексико-морфологического анализа был модифицирован для работы с предметными словарями английского языка, которые удалось совместить со словарями русского языка.

## 5 Анализ существующей схемы разбора

Реализованная ранее схема морфологического анализа выглядела следующим образом.

Сформированы словарь классов окончаний (СКО) — все возможные парадигмы русского языка и СО — основы слов со ссылками на соответствующий класс окончаний. В основе хранится неизменяемая часть слова с постоянными для данного слова морфологическими признаками. Класс окончаний содержит все окончания данной парадигмы плюс изменяемые (т. е. присущие каждому окончанию) морфологические параметры.

Например, слово «бытие» имеет основу «быти» и класс окончаний за номером 1759, содержащий окончания в именительном, родительном, дательном, винительном, творительном и предложном падежах, а именно: «е», «я», «ю», «е», «ем», «и» (множественного числа это слово не имеет). Соответственно в СО имеется запись «быти 1759», а в СКО под номером 1759 закодирована парадигма с указанными окончаниями.

Легко видеть, что в общем случае может иметь место омонимия основ, т. е. несколько одинаковых основ, но с разными парадигмами окончаний. С другой стороны, одна и та же парадигма может «обслуживать» много разных основ. Именно поэтому разделение на СО и СКО оказывается с точки зрения экономии пространства столь эффективным приемом. Однако поиск всех корректных

вариантов разбора требует значительного времени. Действительно, необходимо проверить все варианты разбиения входной лексемы на основу и окончание и для каждого варианта найти в словаре основу и проверить, имеется ли во всех классах окончаний, связанных с этой основой, предполагаемое окончание.

В предыдущей версии был также реализован алгоритм обработки незнакомых системе слов «по аналогии» [4]. В реализации этого метода использовался третий словарь — СХО, в который записываются все 1-, 2-, 3-буквенные и т. д. «хвосты» основ (первые буквы основ отбрасываются) с указанием соответствующего класса окончаний. Было решено, что в СХО не будет одинаковых «хвостов», а его класс окончаний вычисляется из статистических соображений — по максимуму основ в СО, имеющих данный «хвост» и данный класс окончаний. Если слово не находится в СО, то та же схема анализа повторяется, но уже с помощью пары словарей СХО–СКО.

Отметим, что в предыдущей версии все словари были реализованы в виде деревьев, каждый уровень которых соответствовал одной букве (причем порядок в основах и окончаниях был инвертированным). Это давало возможность резко уменьшить рабочий объем словарей. Однако поиск осуществлялся далеко не лучшим в смысле времени работы образом: путем поиска нужной ветки и спуска на следующий уровень по веткам дерева.

Описанную реализацию можно было улучшать, добиваясь определенного прироста быстродействия. Но достичь качественного улучшения представлялось проблематичным.

## **6 Принципы реализации в новой версии**

Прежде всего ставилась задача кардинально повысить скорость анализа русских текстов. Оказалось, что это можно осуществить только путем отказа от принципа разделения на СО и СКО и объединения всей морфологической информации в единый словарь. Разумеется, это резко повышает требования к оперативной памяти. Но в современных компьютерах она уже достигает нескольких гигабайт, что позволяет реализовать этот принцип даже для словаря порядка 100 тыс. слов.

Таким образом, схема информации представляется нарочито упрощенной: единый словарь всех словоформ с полным морфологическим описанием каждой словоформы. Каждая словоформа рассматривается как ключ доступа, по которому производится прямое чтение информации и извлечение морфологических знаний.

Однако, чтобы такая «примитивная» схема реализации была успешной, требуется решить несколько принципиальных задач. Прежде всего, это задача прямого доступа по ключу для больших массивов. В данной реализации предлагается использование принципа хеш-функций, позволяющего реализовать очень быстрый доступ. Принцип работы хеш-функций хорошо известен. Вкратце он сводится к преобразованию ключа в уникальный адрес доступа с помощью ряда

математических и логических операций. Хотя такие вычисления занимают мало времени, однако при реализации этого подхода возникает ряд нетривиальных проблем. Во-первых, это достижение равномерности такого преобразования (в смысле равновероятного покрытия всего адресуемого пространства). Указанная равномерность так или иначе решается разработчиками алгоритма хеширования. Во-вторых, это проблема «коллизий», т. е. разрешения конфликтных ситуаций, когда несколько разных ключей преобразуются в один и тот же адрес. И, наконец, в-третьих, это проблема рационального заполнения «пустот», т. е. элементов массива, остающихся пустыми в силу того, что в них не отображается ни один ключ. Вторые две проблемы необходимо уже решать разработчику системы, использующей хеш-функции.

Помимо корректного использования хеш-функций имеются еще обычные для такого рода систем проблемы ввода и корректировки морфологической информации. Ниже будут описаны методы, позволяющие эти проблемы удовлетворительно решить.

## 7 Структура информационных массивов.

Ради рационального использования памяти вводятся два массива информации: массив адресации (МА) и массив информации (МИ).

Массив адресации представляет собой хеш-массив, в котором располагаются ссылки (смещения) на МА. Этот массив организован в зоны по несколько записей в зоне, причем в 1-м байте зоны хранится число записей. Хеш-функция преобразует ключ (словоформу) в два хеш-кода: Hash1, который указывает на номер зоны, и Hash2 — 3-байтный код, которым идентифицируется запись внутри зоны. Каждая запись включает в себя поле Hash2 и адрес в МИ. Все записи, попадающие в одну и ту же зону (т. е. имеющие один и тот же Hash1), последовательно заполняют эту зону и ищутся в ней по Hash2. При переполнении зоны они пишутся в следующую зону на свободное место.

Двукратное хеширование позволяет достаточно экономно использовать адресуемое пространство, практически не теряя в скорости доступа. Массив адресации управляет несколькими параметрами: величиной массива (в байтах), длиной записи и длиной зоны. Это дает возможность после нескольких попыток подобрать параметры, оптимально подходящие под величину массива морфологической информации.

Массив информации является текстовым массивом, в котором через вычислительный комплекс хранятся (в определенном формате) словоформы и морфологические признаки. Информация в нем структурируется по меточному принципу: информация делится на части, каждая из которых предваряется метками, такими как «*m=*» — меткой морфологических параметров, «*k=*» — меткой семантической информации и т. д. Например:

высадки => | ка *m=52NжXНр* |ка *m=52NжXНви*: |ок *m=52NmХФви*:

Здесь описан морфологический разбор словоформы «высадки», причем даются три варианта разбора: родительный падеж единственного числа слова «высадка»; множественное число того же слова и множественное число существительного «высадок». После символа «|» дается окончание канонической формы, а две цифры после «т=» означают длину основы и длину окончания для разных вариантов разбора слова «высадки».

## **8 Анализ «по аналогии» в новой версии**

В принципе, «прямой» подход к построению морфологического словаря может быть применен и для анализа «по аналогии». Для этого необходимо сгенерировать все усеченные варианты каждой словоформы (с указанием предположительных морфологических признаков), а в процессе обработки входной словоформы, при отсутствии ее в словаре, начать процесс усечения с начала слова с поиском каждого усеченного варианта.

Однако эксперименты показали, что при этом словарь (точнее оба массива: МА и МИ) увеличиваются более чем на порядок, что уже вызывает затруднение при реализации даже на современных персональных компьютерах (тем более, что морфологический анализ — лишь одна из многочисленных задач, решаемых в рамках лингвистического процессора). Поэтому было решено реализовать гибридный подход: слова из словаря распознавать с помощью новой версии, а для анализа «по аналогии» оставить алгоритмы предыдущей версии.

## **9 Корректировка и ввод информации**

Ввод в МА и МИ осуществляется с помощью специального блока загрузки с морфологического массива (ММ). Этот массив имеет вид текстового файла и может корректироваться с помощью любого текстового редактора. Порядок записей в этом файле не имеет значения. Однако требуется, чтобы все варианты морфологического разбора одной словоформы (из-за морфологической омонимии их может быть несколько) были собраны в одну запись. Кроме того, имеется возможность помимо морфологической информации помещать любую другую, относящуюся к данной словоформе лексическую или семантическую информацию. Это дает возможность в перспективе сделать универсальный морфо-семантический словарь. Для объединения разнородной информации в одной записи при формировании файла ММ осуществляется специальный процесс сборки, когда из разных источников в одну запись собирается информация о данной словоформе.

Процесс сборки происходит в несколько стадий с помощью того же блока загрузки, но работающего в режиме «склеивания» с последующей выгрузкой промежуточного варианта МИ в текстовый вид.

## 10 Некоторые параметры новой версии

Общий объем двух массивов (МА и МИ) для словаря в 90 тыс. основ составляет 120 МБ (в старой версии — менее 1 МБ).

Время загрузки словаря — порядка 10 с.

Скорость обработки текста русской деловой прозы увеличилась приблизительно в 4 раза. При этом если учесть, что около 10% слов распознаются «по аналогии», т. е. со старой скоростью, то скорость обработки слов, имеющихся в словаре, увеличилась в 6 раз. Следует отметить, что сам механизм доступа к морфологической информации гораздо более эффективен (по крайней мере в десятки раз). Однако анализ показал, что снижение эффективности происходит из-за того, что все процессы «постморфологии», занимающие достаточно много времени, остались от старой версии.

## 11 Возможности дальнейшего развития

В порядке развития данного проекта намечаются следующие работы:

- перевод всей морфологической системы под ОС UNIX;
- создание новой версии для английской морфологии;
- перенесение некоторых вычислений из фазы «постморфологии» на этап сборки информации при вводе — это позволит еще более повысить эффективность анализа;
- поскольку информация в словаре имеет строго текстовый вид, то появляется возможность реализации на этой основе других компонент лингвистического процессора, например предметных терминологических словарей.

## Литература

1. Сегалович И., Маслов М. Русский морфологический анализ и синтез с генерацией моделей словоизменения для не описанных в словаре слов // Диалог'98: Тр. Междунар. семинара по компьютерной лингвистике и ее приложениям. — Казань: Хэтер, 1998. Т. 2. С. 547–552.
2. Коваленко А. Вероятностный морфологический анализатор русского и украинского языков // Системный администратор, 2002. № 1. <http://www.keva.ru/stemka/stemka.html>.
3. Сокирко А. В. Морфологические модули на сайте [www.aot.ru](http://www.aot.ru) // Компьютерная лингвистика и интеллектуальные технологии: Тр. Междунар. конф. «Диалог-2004». — М.: Наука, 2004. С. 559–564.
4. Сомин Н. В., Соловьева Н. С., Кузнецова Е. В., Шарнин М. М. Система морфологического анализа: опыт эксплуатации и модификации // Системы и средства информатики, 2005. Вып. 15. С. 20–30.
5. Кузнецов И. П. Сомин Н. В. Англо-русская система извлечения знаний из потоков информации в среде Интернет // Системы и средства информатики, 2007. Вып. 17. С. 236–254.

6. Кузнецов И. П., Мацкевич А. Г. Семантико-ориентированные системы на основе баз знаний. — М.: МТУСИ, 2007. 173 с.
7. Сомин Н. В., Кузнецов И. П., Мацкевич А. Г., Николаев В. Г. Методы и средства настройки морфо-лексического анализатора на предметную область // Системы и средства информатики, 2009. Вып. 19. С. 96–118.
8. Кузнецов И. П., Сомин Н. В. Особенности лексико-морфологического анализа при извлечении информационных объектов и связей из текстов естественного языка // Компьютерная лингвистика и интеллектуальные технологии: По мат-лам ежегодной Междунар. конф. «Диалог». — М.: РГГУ, 2010. Вып. 9(16). С. 254–264.
9. Сомин Н. В., Кузнецов И. П., Николаев В. Г., Соловьева Н. С., Мацкевич А. Г. Методы устранения неопределенностей блока лексико-морфологического анализа при извлечении знаний из текстов естественного языка // Системы и средства информатики, 2011. Т. 21. № 2. С. 96–114.
10. Кузнецов И. П., Сомин Н. В., Козеренко Е. Б. Особенности лексико-морфологического анализа в задачах извлечения структур знаний из текстов естественного языка // Искусственный интеллект (журнал НАН Украины), 2011. Т. 3. С. 105–116.
11. Кузнецов И. П., Сомин Н. В. Выявление имплицитной информации из текстов на естественном языке: проблемы и методы // Информатика и её применения, 2012. Т. 6. Вып. 1. С. 49–58.
12. Соловьева Н. С., Сомин Н. В. ТЕРМИН-3 — система динамического гипертекста // Системы и средства информатики, 1995. Вып. 7. С. 95–104.
13. Дягилева А. В., Киселев С. Л., Сомин Н. В. Статистическая модель рубрикации текстов на примере сообщений СМИ // Дистанционное образование, 1998. № 7. С. 16–21.
14. Сомин Н. В., Соловьева Н. С. Рубрицирование текстов как информационная технология // Системы и средства информатики, 2001. Вып. 11. С. 195–201.
15. Сомин Н. В. Каталоги специальных знаний в системе анализа текстов на естественном языке // Проблемы и методы информатики: Тезисы докл. II Научной сессии Института проблем информатики Российской академии наук. — М.: ИПИ РАН, 2005. С. 128–131.

*Поступила в редакцию 12.08.14*

---

## **USING HASH FUNCTION FOR INCREASING SPEED OF WORK OF THE SOFTWARE FOR MORPHOLOGICAL ANALYSIS OF RUSSIAN TEXTS**

*N. V. Somin and M. M. Sharnin*

Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str.,  
Moscow 119333, Russian Federation

**Abstract:** The paper considers the problem of increasing efficiency of morphological analysis of Russian texts. The software system for morphological

analysis is described, including the set of morphological characteristics and the algorithms of work. The paper mentions the software systems solving the problem of logic-semantic analysis of natural language texts in which the software system for morphological analysis found application. Features of the system are discussed from the point of view of occupied memory and work speed. The way of storage of morpholexical information using hash functions is suggested which provides high speed of access. The difficulties arising during realization of such approach are discussed and possible solutions are considered. The paper describes the structure of information arrays of the new version and the search algorithms realized in it. The paper also describes a subsystem for putting in and updating morphological information. Specific parameters of the new realization of the software system for morphological analysis and information on speed of work acceleration in comparison with the previous version are given. The paper discusses opportunities of development of the new version of the software system for morphological analysis and of transferring the suggested approach to other components of the linguistic processor.

**Keywords:** morphological analysis; hash function; linguistic processor; logic-semantic analysis of natural language texts

**DOI:** 10.14357/08696527140315

## References

1. Segalovich, I., and M. Maslov. 1998. Russkiy morfologicheskiy analiz i sintez s generatsiey modeley slovoizmeneniya dlya ne opisannykh v slovare slov [The Russian morphological analysis and synthesis with generation of models of word change for the words not described in the dictionary]. *Conference Dialog'98 Proceedings*. Kazan': Hjeter. 2:547–552.
2. Kovalenko, A. 2002. Veroyatnostnyy morfologicheskiy analizator russkogo i ukrainskogo yazykov [Probabilistic morphological analyzer of the Russian and Ukrainian languages]. *Sistemnyy Administrator* [System Administrator] 1. Available at: <http://www.keva.ru/stemka/stemka.html> (accessed August 12, 2014).
3. Sokirko, A. V. 2004. Morfologicheskie moduli na sayte [www.aot.ru](http://www.aot.ru) [Morphological modules at the site [www.aot.ru](http://www.aot.ru)]. *Conference Dialog'2004 Proceedings*. Verkhnevolzhskiy. Available at: <http://www.aot.ru/docs/sokirko/Dialog2004.htm> (accessed August 12, 2014).
4. Somin, N. V., N. S. Solov'eva, E. V. Kuznetsova, and M. M. Sharnin. 2005. Sistema morfologicheskogo analiza: Opyt ekspluatatsii i modifikatsii [The System of morphological analysis: Operating experience and modifications]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 15:20–30.
5. Kuznetsov, I. P., and N. V. Somin. 2007. Anglo-russkaya sistema izvlecheniya znanii iz potokov informatsii v srede Internet [English-Russian languages system of extraction of knowledge from Internet]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 17:236–254.
6. Kuznetsov, I. P., and A. G. Matskevich. 2007. *Semantic-orientirovannye sistemy na osnove baz znanii* [Semantiko-focused systems on the basis of knowledge bases]. Moscow: MTUCI. 173 p.

7. Somin, N. V., I. P. Kuznetsov, A. G. Matskevich, and V. G. Nikolaev. 2009. Metody i sredstva nastroyki morfo-leksicheskogo analizatora na predmetnuyu oblast' [Methods and means of setup of the morfolexical analyzer for subject domain]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 19:96–118.
8. Kuznetsov, I. P., and N. V. Somin. 2010. Osobennosti leksiko-morfologicheskogo analiza pri izvlechenii informatsionnykh ob'ektov i svyazey iz tekstov estestvennogo jazyka [Features of the lexical-morphological analysis at extraction of information objects and communications from natural language texts]. *Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii. Po mat-lam Mezhdunar. Konf. "Dialog 2010"* [Computational Linguistics and Intellectual Technologies. Conference (International) Dialog 2010 Proceedings]. Moscow: RGGU. 9(16):254–264.
9. Somin, N. V., I. P. Kuznetsov, V. G. Nikolaev, N. S. Solovyeva, and A. G. Matskevich. 2011. Metody ustraneniya neopredelennostey bloka leksiko-morfologicheskogo analiza pri izvlechenii znaniy iz tekstov estestvennogo jazyka [Methods to eliminate uncertainty of lexical-morphological analysis in process of knowledge extraction from natural language texts]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 21(2):96–114.
10. Kuznetsov, I. P., N. V. Somin, and E. B. Kozerenko. 2011. Osobennosti leksiko-morfologicheskogo analiza v zadachakh izvlecheniya struktur znaniy iz tekstov estestvennogo jazyka [Features of the lexical-morphological analysis in problems of extraction of structures of knowledge from natural language texts]. *Iskusstvennyy Intellekt (NAN Ukrayiny)* [Artificial Intelligence (National Academy of Sciences of Ukraine)] 3:105–116.
11. Kuznetsov, I. P., and N. V. Somin. 2012. Vyyavlenie implitsitnoy informatsii iz tekstov na estestvennom jazyke: Problemy i metody [Extraction of implicit information from the natural language texts: Problems and methods]. *Informatika i ee Primeneniya — Inform. Appl.* 6(1):49–58.
12. Solovyeva, N. S., and N. V. Somin. 1995. TERMIN-3 — sistema dinamicheskogo giperteksta [The system of dynamic hypertext]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 7:95–104.
13. Dyagileva, A. V., S. L. Kiselev, and N. V. Somin. 1998. Statisticheskaya model' rubrikatsii tekstov na primere soobshcheniy SMI [The statistical model of a rubrication of texts on mass media examples]. *Distantsionnoe Obrazovanie* [Remote Education] 7:16–21.
14. Somin, N. V., and N. S. Solovyeva. 2001. Rubritsirovaniye tekstov kak informatsionnaya tekhnologiya [Rubrication of texts as information technology]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 11:195–201.
15. Somin, N. V. 2005. Katalogi spetsial'nykh znaniy v sisteme analiza tekstov na estestvennom jazyke [Catalogs of special knowledge in system of the analysis of texts in a natural language]. *Problemy i Metody Informatiki. II Nauchnaya Sessiya IPI RAN. Tezisy dokladov.* [Problems and Methods of Informatics. II Scientific Session of IPI RAN. Abstracts]. Moscow: IPI RAN. 128–131.

Received August 12, 2014

## Contributors

**Somin Nicolay V.** (b. 1947) — Candidate of Science (PhD) in physics and mathematics, leading scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; somin@post.ru

**Sharnin Mihael M.** (b. 1959) — Candidate of Science (PhD) in technology, senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; mc@keywen.com

## К ВОПРОСУ О ЛОГИСТИКЕ ВНЕДРЕНИЯ И ИСПОЛЬЗОВАНИЯ МОБИЛЬНЫХ ЭЛЕКТРОННЫХ УЧЕБНИКОВ

*Д. А. Богданова<sup>1</sup>, А. А. Федосеев<sup>2</sup>*

**Аннотация:** На основе обзора международных и российских публикаций по вопросу внедрения электронных учебников (ЭУ) в школьное образование с учетом существующей ситуации по обеспечению российских школ возможностями широкополосной связи выявлен ряд проблем, сопровождающих этот процесс и в определенном смысле препятствующих его успешному осуществлению и требующих соответствующего решения. Введено понятие идеальной доступности ЭУ, на основе которого предложены способы распространения ЭУ. Рассмотрено понятие жизненного цикла ЭУ, что позволило разграничить процессы модернизации и создания нового ЭУ. Уточнен смысл понятия мобильности ЭУ. В результате удалось сформулировать образ ЭУ, реализация которого позволяет снизить зависимость его внедрения в российские школы от существующей обеспеченности каналами связи. Предложенное решение по разным причинам отличает его от предлагаемых аналогов как в России, так и за рубежом.

**Ключевые слова:** электронный учебник; логистика; жизненный цикл

**DOI:** 10.14357/08696527140316

### 1 Введение

Мобильный ЭУ для школ стал одной из актуальных тем последнего времени. Периодически появляется информация о pilotных проектах по внедрению и использованию школьных ЭУ в разных странах. Так, в 2012 г. Турция приступила к реализации одного из самых масштабных проектов по внедрению планшетов в средних школах. В течение трех лет начиная с 2013 г. 16 млн турецких школьников с 1-го по 12-й класс бесплатно получат от государства планшеты с установленными на них специально разработанными ЭУ, удовлетворяющими требованиям государственной учебной программы. Амбициозные планы поставила перед собой Южная Корея, объявившая о полном переходе к 2015 г. на ЭУ [1]. Одновременно с информацией об успехах появляются материалы о допущенных недоработках и просчетах: не все проекты по внедрению учебников

---

<sup>1</sup>Институт проблем информатики Российской академии наук, d.a.bogdanova@mail.ru

<sup>2</sup>Институт проблем информатики Российской академии наук, A.fedoseev@ipiran.ru

на основе планшетов проходят гладко. Например, осенью 2012 г. в образовательных структурах Лос-Анджелеса разгорелся большой скандал: более трехсот учащихся старших классов, которым в рамках огромного регионального образовательного проекта по созданию ЭУ были выданы планшеты, тут же вскрыли систему безопасности, установленную школой, и отправились кто в Facebook, а кто — общаться в Twitter [2]. При этом, по мнению учащихся, вся система настроек, установленная школой, могла быть легко обойдена любым человеком с минимальным опытом, так же как и школьная система безопасности — Firewall. Образовательные власти региона обвинили в том, что предварительно не было разработано четкой стратегии внедрения и перехода на новые средства обучения, все действия руководства выглядели слишком поспешными и непродуманными. Учащиеся и персонал не были проинструктированы надлежащим образом — и школьники позволили себе подобные вольности.

Стараясь учесть опыт чужих ошибок, Россия, где ЭУ начал постепенно становиться фактом нашей жизни, еще только начинает свой путь. Появляется информация о pilotных проектах по использованию планшетов в отдельных школах отдельных регионов, но происходящее носит скорее единичный, чем массовый характер [1]. В предлагаемой статье авторы представляют свое видение проблемы разработки ЭУ.

## 2 Проблемы, ожидающие решения

Следует обратить внимание на тот факт, что, несмотря на значительное число зарубежных публикаций о внедрении и использовании ЭУ на основе планшетов, многие существенные детали, касающиеся организационных аспектов, не находят освещения, их предварительная тщательная проработка является очень важной. В данном случае имеются в виду не методики обучения и образовательное содержательное наполнение, а вопросы логистики: организация получения, использования и поддержки учебников. В перспективе перехода российских школ к использованию ЭУ имеет смысл проработать эти столь существенные вопросы заранее, чтобы оптимизировать процесс перехода на новые технические средства, а не принимать решения на ходу, уже находясь «в процессе», как это произошло, к примеру, в Лос-Анджелесе или Нью-Йорке (о чем еще будет сказано).

Рассмотрим, какие проблемы, требующие решения, препятствуют немедленному и массовому внедрению ЭУ.

**Начнем с покупки планшета.** Могут использоваться разные схемы:

- государство покупает и отдает детям в безвозмездное пользование, как, например, в Турции, о чем уже было сказано;
- родители покупают за свои деньги по розничной цене. Обычно подобные закупки, с учетом требуемых параметров, координируют активисты родительского комитета, осуществляя их централизованно;

- покупка осуществляется из расчета 50 : 50, где половину стоимости оплачивает государство;
- родители покупают планшет индивидуально, но по оптовой цене, а договор с поставщиком регулируют или школа, или местные органы образования;
- родителям предоставляется низкоконтрольный кредит на покупку. Этот вопрос также регулируется или руководством школы, или местными органами образования.

**Далее возникает проблема обслуживания.** В школе должна возникнуть специальная штатная единица, которая будет заниматься закачиванием новых учебных материалов, поддержкой содержания. Возможно, что этот же сотрудник будет осуществлять и техническое обслуживание планшетов. Эта схема, очевидно, пригодна для ситуации, когда планшет является полной или частичной собственностью школы. В варианте, когда владельцем планшета является семья учащегося, ситуация не столь очевидна. Станет ли школа бесплатно обслуживать планшеты детей или родителям придется платить некоторую сумму за обслуживание? Кроме того, возникает проблема подзарядки аккумуляторов. Каким образом она будет решаться — также не ясно.

**Проблема использования и хранения.** Будут ли учащиеся брать планшет с собой домой или он будет находиться на хранении в школе? В этом случае возникает вопрос хранения: в каком классе, кто будет ответственным за сохранность.

При этом следует помнить о том, что планшетом будут пользоваться каждый день. Значит, в случае хранения в школе не должно возникать проблем с получением каждым учащимся своего планшета перед началом занятий. В случае если планшет берется домой, кто ответствен за его сохранность по дороге в школу и обратно? В США опубликована статистика, иллюстрирующая тот факт, что некоторые учащиеся из малообеспеченных семей, получив бесплатно планшет в свое пользование, вскоре объявляли его утерянным. После нескольких подобных случаев было найдено решение, когда «потерявшему» не выдавался взамен новый планшет, а предлагалось оставаться в школе после уроков для выполнения домашнего задания на школьном планшете [1].

**Проблема возобновления.** Очевидно, что если ребенок будет ходить с планшетом в школу и домой, повышается вероятность его порчи. Или, например, планшет могут у ребенка отнять. Следовательно, возникает проблема страхования планшетов. При этом не ясно, кто будет оплачивать страховой взнос, кто будет выгодоприобретателем и т. д.

Высказываются предложения, что планшет должен стать своего рода «портфолио», в котором будут храниться все результаты тестов, контрольных работ и т. д. В этом случае возникает вопрос о том, каким образом будет создаваться backup (для такого случая он просто необходим) и где он будет находиться, куда результаты будут передаваться и насколько это хранилище будет защищено от внешнего вторжения как физического, так и удаленного, программного.

**Проблема кибербезопасности.** С появлением подхода, реализующего принцип «удаленной, домашней работы сотрудника» очень много вопросов возникло в связи с необходимостью дополнительных мер по обеспечению безопасности локальной сети компании, к которой «удаленный сотрудник» будет подключаться из дома или после работы дома. BYOD — bring your own device — этой аббревиатурой обозначают упомянутую ситуацию. Проблема касается как опасностей вирусного заражения, так и сохранности информации, хранящейся на сервере компании. Аналогичная проблема возникнет и в ситуации, когда учащиеся, прия с планшетами, которыми они пользовались дома, станут подключаться к школьной сети.

**Проблема пропускной способности каналов связи.** Школьные сети и внешние каналы связи как в России, так и за рубежом создавались, исходя из возможностей школы, имеющей несколько компьютерных классов, которые могли выходить в Интернет не одновременно, в зависимости от школьного расписания. Когда все классы будут использовать мобильные учебники во время уроков, может сложиться ситуация, когда возможности каналов связи будут не в состоянии обеспечить требуемой пропускной способности. Так, в 2011–2012 учебном году в ряде школ Нью-Йорка было запрещено использовать мобильные устройства для подключения к школьному Wi-Fi, поскольку массовое подключение iPad, iPod Touch, телефонов iPhone и Android превысило возможности почтового сервера [1]. Пришлось отдельно и в срочном порядке решать проблему повышения пропускной способности каналов связи для нью-йоркских школ. Поэтому неслучайно разработчиком требований к ЭУ в США выступило Министерство связи по заказу Министерства образования и проблема рассматривалась комплексно: с позиций требований и к каналам связи, и к самим устройствам. Ведь концепция идеальной доступности (см. далее), по мнению авторов, помимо онлайнового контента предполагает и возможность постоянного неограниченного доступа, что обеспечивается либо широкополосными каналами связи, либо специальными технологиями, что актуально в нынешних российских условиях, поскольку лишь 10% городских и 3% сельских школ имеют скоростной Интернет [3].

**Работа с клавиатурой.** Переход на ЭУ повышает требования к скорости работы с клавиатурой. В связи с чем необходимо обучение навыкам слепой десятипалцевой печати. В западных школах, в частности в США и Великобритании, существуют методики обучения, разработанные специально для детей с позиций требований санитарных норм относительно положения кисти, нагрузок на пальцы и т. д. Обучение там ведется со 2-го–3-го класса начальной школы. Отрадно отметить, что в России также объявлено обучение десятипалцевой печати в начальной школе. Но в школах, с которыми авторы сотрудничают, эти изменения пока не вошли в практику. И, несмотря на возникшую внешнюю схожесть в этом вопросе с западными школами, здесь, по мнению авторов, кроются определенные сложности. Нельзя не отметить тот факт, что клавиатура (если можно так сказать) на большинстве планшетов отличается от той, для которой были разработаны рекомендованные клавиатурные тренажеры [4–7]. Специали-

сты говорят: «A touch is not a click». Эта фраза, по мнению авторов, отражает принципиальную разницу между сенсорным и клавиатурным письмом.

В случае использования схемы «2 ученика на 1 компьютер» некоторая часть из перечисленных вопросов, в частности BYOD, использования и страхования от порчи, естественным образом снимается, но остальные вопросы, общие со схемой 1 : 1, по-прежнему нуждаются в решении. Нельзя не признать, что, судя по тенденциям развития информационных и коммуникационных технологий (ИКТ), эта схема в ближайшее время, скорее всего, потеряет свою актуальность. Цена ИКТ-устройств имеет тенденцию к снижению, поэтому становится по карману все более широкому кругу пользователей.

Какое решение перечисленных проблем видится возможным и целесообразным?

Как видно из перечисленного выше, существенная часть проблем возникает в связи с наличием связки ЭУ–планшет. Это говорит о том, что ЭУ, соединенный с устройством только определенного типа, не является хорошо продуманным решением. Скорее всего, эта ситуация навязана системе образования производителями электронных устройств и издателей ЭУ. Например, на планшетах компании Apple используются ЭУ, созданные с помощью конструктора ЭУ компании Apple, — и никакие другие. Здесь следует оговориться, что, как правило, какие-либо утверждения, базирующиеся на анализе текущей ситуации в компьютерном мире, очень быстро устаревают. Возможно, к моменту выхода этой статьи ситуация уже не будет столь жесткой. Тем не менее интересно задаться вопросом: а что явились бы хорошо продуманным решением, которое позволило бы избежать большинства из перечисленных проблем. Какой может быть идеальная доступность ЭУ?

### **3 Концепция идеальной доступности электронного учебника**

Понятие идеальной доступности ЭУ заключается в следующем:

- ЭУ доступен ученику с любого компьютера, поддерживающего мультимедийное содержание и интерактивные действия, в том числе дома и в школе. Акцент здесь делается не на том, чтобы разработчик указал, какой тип компьютеров может быть применен, а на том, чтобы обеспечить применимость ЭУ ко всем типам имеющихся на рынке компьютеров, принципиально поддерживающих мультимедийность и интерактивность;
- ученик не носит с собой никаких устройств, чтобы получать доступ к ЭУ;
- где бы и когда бы ученик ни получил доступ к ЭУ, последний идентифицирует ученика и сразу выводит его на незаконченные действия (в соответствии с индивидуальной траекторией обучения), если таковые есть;
- где бы ученик ни работал с ЭУ, учитель с любого компьютера (школьного или домашнего) имеет доступ к результатам его интерактивных действий по выполнению заданий и по индивидуальной траектории обучения.

Поясним эти положения.

Требования мультимедийности и интерактивности отмечаются во многих теоретических и рекомендательных работах (см., например, [8, 9]). Но интерактивность и особенно мультимедийность реализуются на разных вычислительных платформах по-разному. Поэтому производители и привязывают ЭУ к конкретным типам компьютеров или планшетов. Идеальное представление в данном случае означает, что ЭУ должен работать на вычислительном устройстве любого типа из имеющихся на рынке и принципиально поддерживающих мультимедийность и интерактивность.

Если предполагается, что для обеспечения нормального учебного процесса учащиеся должны ежедневно приносить с собой в школу какие-либо устройства, то школа попадает в зависимость от выполнения этого требования. Устройство может быть забыто дома, его могут отнять по дороге в школу или из школы (дорогостоящие планшеты особенно привлекательны для хищений). Идеальность в данном случае заключается в том, что учащийся работает с тем компьютером, ноутбуком или планшетом, который ему предоставили для работы в школе, дома, в библиотеке или где-либо еще. И при этом ему достаточно знать, как загрузить свой учебник, чтобы начать работу или продолжить ее с того места, на котором был прекращен предыдущий рабочий сеанс.

Что касается фиксации интерактивных действий учащихся и доступа к ним учителей, это определяется требованиями учебного процесса: учителю необходимо знать уровень восприятия учебного материала всеми его учениками [10]. При использовании полиграфического учебника достижение такого результата невозможно. Учитель получает информацию о степени восприятия материала путем опроса нескольких учеников класса. Как правило, один и тот же ученик может быть опрошен учителем несколько раз в течение отчетного периода (четверти или триместра). Учитель более-менее равномерно опрашивает учеников класса в течение (пусть) четверти, с тем чтобы иметь основание выставить оценку каждому ученику в конце периода. Такая практика не позволяет опросить всех учеников по каждой изучаемой теме, поэтому текущий контроль переносится на другие формы: контрольную работу, экзамен и т. п. Как известно, ученики пользуются своими предположениями относительно того, кто стоит на очереди для опроса учителем, и не учат заданный урок, например, после успешного ответа на предыдущем уроке, что заставляет учителя хитрить, делать провокационные опросы недавно опрошенных учеников, но при этом не опрашиваются те ученики, которые должны быть опрошены по плану. В общем, ситуация мало продуктивная и слабо поддающаяся улучшению.

Использование ЭУ принципиально позволяет получать данные об уровне восприятия учебного материала всеми учениками по каждому уроку, что позволяет ликвидировать отмеченный недостаток, если только конструкция ЭУ позволяет это осуществить. Именно в этом, по мнению авторов, заключается тот новый функционал, который откроет ЭУ дорогу в школу [11]. Если новые

технологии (в данном случае в виде ЭУ) ничего нового в учебный процесс не привносят, то пропадает смысл их применения.

#### **4 О возможности реализации идеальной доступности для электронного учебника**

Прежде всего следует отметить, что указанные выше свойства идеальной доступности ЭУ не могут быть реализованы вне сети Интернет. Требование доступа к ЭУ, причем к той его части, на которой ученик остановился во время предыдущего сеанса работы с учебником, возможно, на другом компьютере, однозначно соответствует хранению соответствующих данных в сети Интернет. То же относится и к требованию доступа учителя к результатам работы ученика, осуществляемого с любого компьютера. Таким образом, как управление процессом обучения, со стороны ЭУ, так и подготовка данных для контроля учителем, осуществляются в сети Интернет. Более всего описанному процессу соответствует технология SaaS (Software as a Service).

Значительно усложняющим ситуацию обстоятельством является тот факт, что в настоящее время не существует возможности получить единый программный код ЭУ, который правильно и полноценно отображал бы задуманные производителем мультимедийные возможности на всех существующих платформах компьютеров и планшетов. Эта ситуация требует от производителей разработки своего программного кода ЭУ для каждой платформы.

Размещение всего ЭУ (так же как и комплекта учебников для каждого класса) в сети Интернет приведет к необходимости неоправданной массивной перекачки информации, что существенно осложнит или сделает невозможной нормальную работу в большинстве школ. Правильнее будет, если постоянная, не подвергающаяся частым изменениям мультимедийная информация будет содержаться на компьютерах, с которыми учащимся предполагается работать. И только переменная, довольно скромная часть информации будет путешествовать в Интернете. Сюда относятся такие данные, как идентификация учащегося, в том числе как ученика определенного класса определенной школы, место его оставновки в изучаемом материале, маршрут продвижения по содержанию учебника, итоги выполнения заданий и некоторые другие служебные данные. Такой способ разделения данных ЭУ более соответствует технологии S+S.

Электронный учебник, размещенный в сети Интернет или разделенный описанным образом, отчасти снимает проблему пиратства, что является дополнительным преимуществом этого метода. Размещаемые на школьных и домашних компьютерах части ЭУ не могут использоваться в учебном процессе предусмотренным разработчиком способом без информации, хранящейся в сети Интернет, например на сайте производителя, что делает бессмысленным их несанкционированное копирование. Более того, эти части могут распространяться бесплатно, по цене носителя и размещаться на стольких компьютерах, сколько необходимо для того, чтобы каждый учащийся имел полноценный и повсеместный доступ к своим

ЭУ в любое удобное ему время. Платной может быть только организующая учебный процесс часть ЭУ, хранящаяся в сети Интернет.

Доступ к ЭУ с любого компьютера (естественно, обладающего необходимыми свойствами воспроизведения мультимедийных файлов и интерактивных алгоритмов) решает также ряд проблем, связанных с невозможностью по каким-либо причинам установки ЭУ на некоторые конкретные устройства либо с порчей этих устройств. Поскольку необходимая для обеспечения учебного процесса информация хранится не на отдельном устройстве, а в сети Интернет, достаточно воспользоваться альтернативным устройством на время починки или замены основного.

Очевидно, что учебное заведение не несет ответственности за выбор компьютеров или планшетов, которые покупаются семьями учеников. Тем не менее учебное учреждение может потребовать соблюдения некоторых условий для работы учащихся с ЭУ вне школы. Это относится к соблюдению санитарно-гигиенических норм, таких как размер и разрешение экрана, а также приемлемый тип клавиатуры и других необходимых внешних устройств. При невозможности соблюдения требований, например по финансовым соображениям или из-за отсутствия требуемых устройств в продаже, учебное заведение должно предоставить необходимое количество соответствующих требований компьютеров для работы учащихся в школе во внеурочное время.

## 5 Жизненный цикл электронного учебника

Электронный учебник является изделием (продуктом) и, как любое другое изделие, должен обладать некоторыми характеристиками, свойственными изделию. Одной из основных характеристик изделий является их жизненный цикл. Вместе с тем, поскольку ЭУ является информационным изделием, у него есть специфические черты по сравнению с техническими изделиями. Однажды созданное информационное изделие может затем копироваться и продаваться неограниченное число раз без каких-либо затрат на производство (за исключением процессов копирования и упаковки). Тем не менее существуют затраты на обслуживание уже проданных ЭУ. Для того чтобы разобраться в этом вопросе, следует сформулировать, что представляет собой ЭУ как изделие.

Сутью и смыслом учебника является система использованных в нем педагогических решений совместно с контентом учебника в виде текста, иллюстраций, заданий и т. п. Электронный учебник несет мультимедийный и интерактивный контент. Таким образом, ЭУ как изделие образуется комплексом педагогических решений, мультимедийным и интерактивным контентом, системой измерения и оценки уровня воспринимаемого материала, реализованных в виде программного кода для некоторой вычислительной платформы. Программная реализация этого ЭУ для иной вычислительной платформы является тем же самым изделием, предназначенным для использования на этой платформе, хотя программный код отличается от первоначального.

Электронный учебник может и должен быть модернизирован в следующих случаях:

- исправление замеченных недостатков или ошибок в контенте;
- изменение стандарта обязательного содержания предмета: некоторые разделы могут быть убраны или добавлены в образовательную программу;
- внедрение в общепринятую практику и добавление в ЭУ новых, ранее не реализуемых способов взаимодействия ЭУ с учеником (новая мультимедийность и интерактивность).

Указанные изменения приводят к модернизации ЭУ. Если произошла модернизация, то прежний ЭУ должен быть заменен на модернизированный для всех пользователей без каких-либо доплат или условий.

Если же изменяются педагогические основы ЭУ, а также система мультимедийного и интерактивного контента, то это приводит к появлению другого учебника, т. е. к появлению нового изделия. При этом прежний ЭУ, уже приобретенный образовательными учреждениями, может продолжать использоваться до тех пор, пока не будут приняты решения относительно перехода на другой ЭУ. Таким образом, жизненный цикл ЭУ продолжается до тех пор, пока его использование не прекратится последним пользователем (учебным заведением). Здесь следует иметь в виду, что учебник используется учеником не только во время изучения предмета, но также и после, при повторении, во время подготовки к экзаменам. Вообще, пока ученик учится в школе и далее вплоть до поступления в вуз, он должен иметь доступ к своему ЭУ. А может быть, следует сохранять доступ на протяжении всей трудовой жизни ученика?

Программная реализация ЭУ для определенной платформы (или кроссплатформенная) может существовать до тех пор, пока существует платформа или комплекс платформ. Изменение программной реализации, сохраняющей изделие — ЭУ:

- изменение программной среды (системы) конструктора учебников;
- добавление новой платформы.

Переход на новый компьютер или на новую платформу (на которой ЭУ не был ранее реализован) не означает возникновения нового изделия. Продукт существует, пока реализуется на вновь возникающих платформах. Жизненный цикл прекращается, если ЭУ не поддерживается на вновь возникающих и заполняющих рынок компьютерных платформах, а прежние платформы выходят из употребления и не поддерживаются производителями.

## **6 Заключение**

В первой части статьи были перечислены некоторые проблемы, от решения которых зависят способы внедрения ЭУ в повседневную школьную жизнь. Как эти проблемы выглядят с позиций предложенного в статье подхода? Очевидно,

что все проблемы, связанные с обеспечением учеников планшетами, полностью уходят. Исчезают проблемы обслуживания, хранения, возобновления, кибербезопасности. Единственная проблема — проблема пропускной способности школьных каналов связи — оказывается решенной не полностью. Тем не менее предложен метод, существенно снижающий требования к пропускной способности каналов.

Если признать, что мобильность ЭУ заключается не в том, чтобы таскать его с собой, а в том, чтобы иметь доступ к ЭУ там, где это предпочтительно по тем или иным причинам, то окажется, что большинство проблем сами собой исчезают.

На основе рассмотрения зарубежного опыта распространения ЭУ и создания идеального представления относительно доступности ЭУ, авторы сформулировали предложение по логистике ЭУ, свободное от большинства приведенных проблем. Решение, освобождающее ЭУ от конкретного носителя в виде планшета или компьютера определенного типа, с одной стороны, позволяет достичь желательной гибкости в практическом применении ЭУ. С другой стороны, оно несколько осложняет жизнь производителей ЭУ. Однако следует ожидать, что, во-первых, выгоды от универсальности применения ЭУ превысят затраты по его реализации для всех существующих платформ, а во-вторых, развитие программирования не остановилось и следует ожидать появления универсальных языков, адекватно работающих на платформах всех типов. В любом случае исходить следует из требований, предъявляемых особенностями учебного процесса, а не из диктата производителей вычислительных устройств и ЭУ.

В мире информационных технологий все меняется очень быстро. Поэтому вполне возможно, что в ближайшем будущем и до России дойдут изменения, в настоящее время происходящие в сфере информационных технологий в американских школах, отказывающихся от использования iPad-ов, а компании, занимавшие лидирующие позиции в области разработок для образования, начинают уступать свое лидерство.

## Литература

1. Богданова Д. А. Об электронном учебнике // Дистанционное и виртуальное обучение, 2013. № 5. С. 18–27.
2. Rogers J. School district's iPad plan backfires. <http://www.timesleader.com/news/national-news/886259/School-districts-iPad-plan-backfires>.
3. Медведев призвал увеличить скорость школьного Интернета // CNEWS, 21.05.2014. <http://www.comnews.ru/node/85533>.
4. ФГОС НОО: Учебно-методический портал. [nachalka.seminfo.ru](http://nachalka.seminfo.ru).
5. Программа курса «Информатика и ИКТ». <http://informatika-rf.webnode.ru/programma-kursa-informatika-i-ikt>.
6. Босова Л. Л., Босова А. Ю. Учебно-методический комплект по информатике для 5–7 классов // Лаборатория знаний, 2012. Вып. 9. <http://gazeta.lbz.ru/2012/9/9nomer.pdf>.

7. Угринович Н. Д. Информатика и ИКТ: Учебник для 7 класса // Образовательные ресурсы Интернета. <http://www.alleng.ru/d/comp/comp163.htm>.
8. Осин А. В. Электронные образовательные ресурсы нового поколения: открытые образовательные модульные мультимедиа системы // Интернет-порталы: содержание и технологии. — М.: Просвещение, 2007. Вып. 4. С. 12–29.
9. Электронные учебники: рекомендации по разработке. — М.: Федеральный институт развития образования, 2012. 24 с.
10. Писарев В. Е., Писарева Т. Е. Теория педагогики. — Воронеж: Квarta, 2009. 611 с.
11. Федосеев А. А. Смешанное обучение: пристальный взгляд // Дистанционное и виртуальное обучение, 2013. № 10. С. 115–125.

Поступила в редакцию 27.06.14

---

## ON LOGISTICS OF MOBILE ELECTRONIC TEXTBOOKS IMPLEMENTATION AND USE

**D. A. Bogdanova and A. A. Fedoseev**

Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str.,  
Moscow 119333, Russian Federation

**Abstract:** Taking into consideration the current situation of supply of broadband access in the schools in Russia, based on reviews in international and national publications regarding introduction of electronic textbooks, the authors identify a number of issues inherent to this process, inhibiting its successful implementation and needing a proper resolution. The authors introduce the concept of ideal availability of an electronic textbook, based on which the authors suggest a number of ways of distribution of electronic textbooks. The authors introduce the notion of lifecycle of an electronic textbook, which helps to identify the processes of modernization and of creation of new textbooks. The authors also clarify the meaning of the notion of mobility of an electronic textbook. As a result, the authors managed to coin a prototype of an electronic textbook, which, when realized, allows minimizing its dependence on availability of broadband access. Therefore, this solution differs from its analogues both in Russia and globally.

**Keywords:** electronic textbooks; logistics; life cycle

**DOI:** 10.14357/08696527140316

## References

1. Bogdanova, D. A. 2013. Ob elektronnom uchebnike [On electronic textbook]. *Distantionnoe i Virtual'noe Obuchenie* [Distance and Virtual Learning] 5:18–27.
2. Rogers, J. 2013. School district's iPad plan backfires. Available at: <http://www.timesleader.com/news/national-news/886259/School-districts-iPad-plan-backfires> (accessed April 6, 2014).

3. Medvedev prizval uvelichit' skorost' shkol'nogo Interneta [Medvedev called for increasing rate of school Internet]. May 21, 2014. *CNEWS*. Available at: <http://www.comnews.ru/node/85533> (accessed August 26, 2014).
4. FGOS NOO [Federal State Educational Standard of primary education]. *Uchebno-Metodicheskiy Portal* [Educational-Methodical Portal]. Available at: <http://nachalka.seminfo.ru> (accessed August 26, 2014).
5. Programma kursa "Informatika i IKT" [Program of the course "Informatics and ICT"]. 2014. Available at: <http://informatika-rf.webnode.ru/programma-kursa-informatika-i-ikt/> (accessed August 26, 2014).
6. Bosova, L. L., and A. Yu. Bosova. 2012. Uchebno-metodicheskiy komplekt po informatike dlya 5–7 klassov [Teaching kit in informatics for 5–7 classes]. *Gazeta "Laboratoriya Znaniy"* [The newspaper "Laboratory of Knowledge"] 9. Available at: <http://gazeta.lbz.ru/2012/9/9nomer.pdf> (accessed August 26, 2014).
7. Ugrinovich, N. D. 2010. Informatika i IKT. Uchebnik dlya 7 klassa [Informatics and ICT. Textbook for grade 7]. *Obrazovatel'nye Resursy Internet* [Educational Resources of the Internet]. Available at: <http://www.alleng.ru/d/comp/comp163.htm> (accessed August 26, 2014).
8. Osin, A. V. 2007. Elektronnye obrazovatel'nye resursy novogo pokoleniya: otkrytye obrazovatel'nye modul'nye mul'timedia sistemy [E-learning resources of the new generation: Modular open educational multimedia systems]. *Internet-portaly: Soderzhanie i tekhnologii* [Internet portals: Content and technology]. Moscow: Prosvetshchenie Publs. 4:12–29.
9. *Elektronnye uchebniki: rekomendatsii po razrabotke* [Electronic textbooks: Recommendations for development]. 2012. Moscow: Federal'nyy institut razvitiya obrazovaniya [Federal Institute for Education Development]. 24 p.
10. Pisarev, E. V., and T. E. Pisareva. 2009. *Teoriya pedagogiki* [Pedagogy theory]. Voronezh: Quarta Publs. 611 p.
11. Fedoseev, A. A. 2013. Smeshannoe obuchenie: Pristal'nyy vzglyad [Blended learning: A closer look]. *Distantsionnoe i Virtual'noe Obuchenie* [Distance and Virtual Learning] 10:115–125.

Received June 27, 2014

## Contributors

**Bogdanova Diana A.** (b. 1946) — Candidate of Science (PhD) in education, senior scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; [d.a.bogdanova@mail.ru](mailto:d.a.bogdanova@mail.ru)

**Fedoseev Andrey A.** (b. 1946) — Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problems, Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; [A.fedoseev@ipiran.ru](mailto:A.fedoseev@ipiran.ru)

---

## ОБ АВТОРАХ

---

**Белоусов Василий Владимирович** (р. 1977) — кандидат технических наук, заведующий сектором Института проблем информатики Российской академии наук

**Бобков Сергей Геннадьевич** (р. 1955) — доктор технических наук, заведующий отделением Научно-исследовательского института системных исследований Российской академии наук

**Богданова Диана Александровна** (р. 1946) — кандидат педагогических наук, старший научный сотрудник Института проблем информатики Российской академии наук

**Бондаренко Ольга Алексеевна** (р. 1970) — младший научный сотрудник Института проблем информатики Российской академии наук

**Волович Константин Иосифович** (р. 1970) — старший научный сотрудник Института проблем информатики Российской академии наук

**Гаранин Александр Иванович** (р. 1951) — кандидат технических наук, старший научный сотрудник Института проблем информатики Российской академии наук

**Грушо Александр Александрович** (р. 1946) — доктор физико-математических наук, член-корреспондент Академии криптографии РФ; ведущий научный сотрудник Института проблем информатики Российской академии наук; профессор факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

**Грушо Николай Александрович** (р. 1982) — кандидат физико-математических наук, старший научный сотрудник Института проблем информатики Российской академии наук

**Дьяченко Юрий Георгиевич** (р. 1958) — кандидат технических наук, старший научный сотрудник Института проблем информатики Российской академии наук

**Егоров Владимир Борисович** (р. 1948) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Российской академии наук

**Зацаринный Александр Алексеевич** (р. 1951) — доктор технических наук, профессор, заместитель директора Института проблем информатики Российской академии наук

**Ильин Александр Владимирович** (р. 1975) — кандидат технических наук, старший научный сотрудник Института проблем информатики Российской академии наук

**Ильин Владимир Дмитриевич** (р. 1937) — доктор технических наук, профессор, заведующий лабораторией Института проблем информатики Российской академии наук

- Ионенков Юрий Сергеевич** (р. 1956) — старший научный сотрудник Института проблем информатики Российской академии наук
- Кириков Игорь Александрович** (р. 1955) — кандидат технических наук, директор Калининградского филиала Института проблем информатики Российской академии наук
- Козлов Сергей Витальевич** (р. 1955) — кандидат технических наук, заведующий отделом Института проблем информатики Российской академии наук
- Колесников Александр Васильевич** (р. 1948) — доктор технических наук, профессор Балтийского федерального университета им. Иммануила Канта; старший научный сотрудник Калининградского филиала Института проблем информатики Российской академии наук
- Кондрашев Вадим Адольфович** (р. 1963) — старший научный сотрудник Института проблем информатики Российской академии наук
- Корепанов Эдуард Рудольфович** (р. 1966) — кандидат технических наук, заведующий сектором Института проблем информатики Российской академии наук
- Королёв Сергей Павлович** (р. 1985) — научный сотрудник Вычислительного центра Дальневосточного отделения Российской академии наук
- Кульберг Николай Сергеевич** (р. 1970) — кандидат физико-математических наук, старший научный сотрудник Вычислительного центра им. А. А. Дородницына Российской академии наук
- Морозов Николай Викторович** (р. 1956) — старший научный сотрудник Института проблем информатики Российской академии наук
- Рождественскене Аста Винценто** (р. 1964) — старший научный сотрудник Института проблем информатики Российской академии наук
- Рождественский Юрий Владимирович** (р. 1952) — кандидат технических наук, заведующий сектором Института проблем информатики Российской академии наук
- Румовская София Борисовна** (р. 1985) — программист 1-й категории Калининградского филиала Института проблем информатики Российской академии наук
- Сергеев Игорь Викторович** (р. 1965) — кандидат технических наук, заместитель директора Института проблем информатики Российской академии наук
- Синицын Владимир Игоревич** (р. 1968) — доктор физико-математических наук, доцент, заведующий отделом Института проблем информатики Российской академии наук
- Синицын Игорь Николаевич** (р. 1940) — доктор технических наук, профессор, заслуженный деятель науки РФ, заведующий отделом Института проблем информатики Российской академии наук
- Соколов Игорь Анатольевич** (р. 1954) — академик Российской академии наук, доктор технических наук, директор Института проблем информатики Российской академии наук

**Сомин Николай Владимирович** (р. 1947) — кандидат физико-математических наук, ведущий научный сотрудник Института проблем информатики Российской академии наук

**Сорокин Алексей Анатольевич** (р. 1980) — кандидат технических наук, заведующий лабораторией Вычислительного центра Дальневосточного отделения Российской академии наук

**Степченков Дмитрий Юрьевич** (р. 1973) — старший научный сотрудник Института проблем информатики Российской академии наук

**Степченков Юрий Афанасьевич** (р. 1951) — кандидат технических наук, заведующий отделом Института проблем информатики Российской академии наук

**Сурков Алексей Викторович** (р. 1978) — старший научный сотрудник Научно-исследовательского института системных исследований Российской академии наук

**Тарасов Алексей Григорьевич** (р. 1982) — кандидат технических наук, старший научный сотрудник Вычислительного центра Дальневосточного отделения Российской академии наук

**Тимонина Елена Евгеньевна** (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Российской академии наук

**Федосеев Андрей Алексеевич** (р. 1946) — кандидат технических наук, ведущий

научный сотрудник Института проблем информатики Российской академии наук

**Хазиахметов Максим Шамилевич** (р. 1989) — аспирант факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

**Шарнин Михаил Михайлович** (р. 1959) — кандидат технических наук, старший научный сотрудник Института проблем информатики Российской академии наук

**Шоргин Всеволод Сергеевич** (р. 1978) — кандидат технических наук, старший научный сотрудник Института проблем информатики Российской академии наук

**Шоргин Сергей Яковлевич** (р. 1952) — доктор физико-математических наук, профессор, заместитель директора Института проблем информатики Российской академии наук

**Яковleva Татьяна Викторовна** (р. 1956) — кандидат физико-математических наук, старший научный сотрудник Вычислительного центра им. А. А. Дородницына Российской академии наук

## **Правила подготовки рукописей статей для публикации в журнале «Системы и средства информатики»**

Журнал «Системы и средства информатики» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информационных технологий.

Журнал издается на русском языке. По специальному решению редколлегии отдельные статьи могут печататься на английском языке.

Тематика журнала охватывает следующие направления:

- информационно-телекоммуникационные системы и средства их построения;
- архитектура и программное обеспечение вычислительных машин, комплексов и сетей;
- методы и средства защиты информации.

1. В журнале печатаются статьи, содержащие результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях.

Публикация не должна нарушать закон об авторских правах.

Направляя рукопись в редакцию, авторы сохраняют все права собственников данной рукописи и при этом передают учредителям и редколлегии неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на ее распространение в России и за рубежом. Авторы должны предоставить в редакцию письмо в следующей форме:

### ***Соглашение о передаче права на публикацию:***

*«Мы, нижеподписавшиеся, авторы рукописи «. . .», передаем учредителям и редколлегии журнала «Системы и средства информатики» неисключительное право опубликовать данную рукопись статьи на русском языке как в печатной, так и в электронной версиях журнала. Мы подтверждаем, что данная публикация не нарушает авторских прав других лиц или организаций.*

*Подписи авторов: (ф. и. о., дата, адрес)».*

Это соглашение может быть предоставлено в бумажном виде или в виде отсканированной копии (с подписями авторов).

Редколлегия вправе запросить у авторов экспертное заключение о возможности публикации представленной статьи в открытой печати.

2. К статье прилагаются данные автора (авторов) (см. п. 8). При наличии нескольких авторов указывается фамилия автора, ответственного за переписку с редакцией.

3. Редакция журнала осуществляет экспертизу присланных статей в соответствии с принятой в журнале процедурой рецензирования.

Возвращение рукописи на доработку не означает ее принятия к печати.

Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редколлегии о публикации статьи или ее отклонении сообщается авторам.

Редколлегия может также направить авторам текст рецензии на их статью. Дискуссия по поводу отклоненных статей не ведется.

5. Редактура статей высылается авторам для просмотра. Замечания к редактуре должны быть присланы авторами в кратчайшие сроки.

6. Рукопись предоставляется в электронном виде в форматах MS WORD (.doc или .docx) или L<sup>A</sup>T<sub>E</sub>X (.tex), дополнительно — в формате .pdf, на дискете, лазерном диске или электронной почтой. Предоставление бумажной рукописи необязательно.

7. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки.

Параметры страницы: формат — А4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху и снизу — 2, от края до нижнего колонтитула — 1,3.

Основной текст: стиль — «Обычный», шрифт — Times New Roman, размер — 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине.

Рекомендуемый объем рукописи — не свыше 20 страниц указанного формата.

Сокращения слов, помимо стандартных, не допускаются. Допускается минимальное количество аббревиатур.

Все страницы рукописи нумеруются.

Шаблоны примеров оформления представлены в Интернете:

<http://www.ipiran.ru/publications/collected/template.doc>

8. Статья должна содержать следующую информацию на **русском и английском языках**:

- Название статьи.
- Ф.И.О. авторов, на английском можно только имя и фамилию.
- Место работы, с указанием города и страны и электронного адреса каждого автора.
- Сведения об авторах, в соответствии с форматом, образцы которого представлены на страницах:  
[http://www.ipiran.ru/journal/collected/2012\\_22\\_02\\_rus/authors.asp](http://www.ipiran.ru/journal/collected/2012_22_02_rus/authors.asp) и  
[http://www.ipiran.ru/journal/collected/2012\\_22\\_02\\_eng/authors.asp](http://www.ipiran.ru/journal/collected/2012_22_02_eng/authors.asp)
- Аннотация (не менее 100 слов на каждом из языков). Аннотация — это краткое резюме работы, которое может публиковаться отдельно. Она является основным источником информации в информационных системах и базах данных. Английская аннотация должна быть оригинальной, может не быть дословным переводом русского текста и должна быть написана хорошим английским языком. В аннотации не должно быть ссылок на литературу и, по возможности, формул.
- Ключевые слова — желательно из принятых в мировой научно-технической литературе тематических тезаурусов. Предложения не могут быть ключевыми словами.

9. Требования к спискам литературы.

Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в каждом из списков литературы в порядке первых упоминаний.

Списки литературы представляются в двух вариантах:

- (1) **Список литературы к русскоязычной части.** Русские и английские работы — на языке и в алфавите оригинала.
- (2) **References.** Русские работы и работы на других языках — в латинской транслитерации с переводом на английский язык; английские работы и работы на других языках — на языке оригинала.

Необходимо для составления списка “References” пользоваться размещенной на сайте <http://www.translit.ru> бесплатной программой транслитерации русского текста в латиницу, при этом в закладке «варианты. . . » следует выбрать опцию BNG.

Список литературы “References” приводится полностью отдельным блоком, повторяя все позиции из списка литературы к русскоязычной части, независимо от того, имеются или нет в нем иностранные источники. Если в списке литературы к русскоязычной части есть ссылки на иностранные публикации, набранные латиницей, они полностью повторяются в списке “References”.

Примеры ссылок на различные виды публикаций в списке “References”:

**Описание статьи из журнала:**

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

**Описание статьи из электронного журнала:**

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

**Описание материалов конференций:**

Usmanov, T.S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primenением gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma “Novye resursosberegayushchie tekhnologii nedropol’zovaniya i povышeniya neftegazootdachi”* [6th Symposium (International) “New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact” Proceedings]. Moscow. 267–272.

**Описание книги (монографии, сборники):**

Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

**Описание переводной книги** (в списке литературы к русскоязычной части необходимо указать: / Пер. с англ. — после названия книги, а в конце ссылки указать оригинал книги в круглых скобках):

1. В русскоязычной части:

Timoshenko S. P., Young D. H., Weaver W. Колебания в инженерном деле / Пер. с англ. — М.: Машиностроение, 1985. 472 с. (Timoshenko S. P., Young D. H., Weaver W. Vibration problems in engineering. — 4th ed. — N.Y.: Wiley, 1974. 521 p.)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. N.Y.: Wiley. 521 p.

**Описание неопубликованного документа:**

Latypov, A. R., M. M. Khasanov, and V. A. Baikov. 2004. Geology and production (NGT GiD). Certificate on official registration of the computer program No. 2004611198. (In Russian, unpubl.)

**Описание интернет-ресурса:**

Pravila tsitirovaniya istochnikov [Rules for the citing of sources]. Available at: <http://www.scribd.com/doc/1034528/> (accessed February 7, 2011).

**Описание диссертации или автореферата диссертации:**

Semenov, V. I. 2003. Matematicheskoe modelirovanie plazmy v sisteme kompaktnyy tor [Mathematical modeling of the plasma in the compact torus]. D.Sc. Diss. Moscow. 272 p.

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticeskogo slovarya informacionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

**Описание ГОСТа:**

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow: Standardinform Publs. 10 p.

**Описание патента:**

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoj samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

10. Присланные в редакцию материалы авторам не возвращаются.
11. При отправке файлов по электронной почте просим придерживаться следующих правил:
  - указывать в поле subject (тема) название журнала и фамилию автора;
  - использовать attach (присоединение);
  - в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.
12. Журнал «Системы и средства информатики» является некоммерческим изданием. Плата за публикацию не взимается, гонорар авторам не выплачивается.

**Адрес редакции журнала «Системы и средства информатики»:**

Москва 119333, ул. Вавилова, д. 44, корп. 2, ИПИ РАН

Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05

e-mail: [rust@ipiran.ru](mailto:rust@ipiran.ru) (Сейфуль-Мулюков Рустем Бадриевич)

<http://www.ipiran.ru/journal/collected>

## **Requirements for manuscripts submitted to Journal “Systems and Means of Informatics”**

Journal “Systems and Means of Informatics” publishes theoretical, review, and discussion articles on the research and development in the field of information technology.

The journal is published in Russian. By a special decision of the editorial board, some articles can be published in English.

Topics covered include the following areas:

- information and communication systems and tools of their design;
- architecture and software of computational complexes and networks; and
- methods and tools of information protection.

1. The Journal publishes original articles which have not been published before and are not intended for publication in other editions. An article submitted to the Journal must not violate the Copyright law. Sending the manuscript to the Editorial Board, the authors retain all rights of the owners of the manuscript and transfer the nonexclusive rights to publish the article in Russian (or the language of the article, if not Russian) and its distribution in Russia and abroad to the Founders and the Editorial Board. Authors should submit a letter to the Editorial Board in the following form:

### ***Agreement on the transfer of rights to publish:***

*“We, the undersigned authors of the manuscript “. . . ,” pass to the Founder and the Editorial Board of the Journal “Systems and Means of Informatics” the nonexclusive right to publish the manuscript of the article in Russian (or in English) in both print and electronic versions of the Journal. We affirm that this publication does not violate the Copyright of other persons or organizations.*

*Author(s) signature(s): (name(s), address(es), date).*”

This agreement should be submitted in paper form or in the form of a scanned copy (signed by the authors).

The Editorial Board has the right to request from the authors an official expert conclusion that the submitted article has no secret data prohibited for publication.

2. A submitted article should be attached with **the data on the author(s)** (see item 8). If there are several authors, the contact person should be indicated who is responsible for correspondence with the Editorial Board and other authors about revisions and final approval of the proofs.
3. The Editorial Board of the Journal examines the article according to the established reviewing procedure. If authors receive their article for correction after reviewing, it does not mean that the article is approved to be published. The corrected article should be sent to the Editorial Board for the subsequent review and approval.
4. The decision on the article publication or its rejection is communicated to the authors. The Editorial Board may also send the reviews on the submitted articles to the authors. Any discussion upon the rejected articles is not possible.
5. The edited articles will be sent to the authors for proofread. The comments of the authors to the edited text of the article should be sent to the Editorial Board as soon as possible.
6. The manuscript of the article should be presented electronically in the MS WORD (.doc or .docx) or L<sup>A</sup>T<sub>E</sub>X (.tex) formats, and additionally in the .pdf format. All documents may be sent by e-mail or provided on a CD or diskette. A hard copy submission is not necessary.

7. The recommended typesetting instructions for manuscript.

Pages parameters: format A4, portrait orientation, document margins (cm): left — 2.5, right — 1.5, above — 2.0, below — 2.0, footer 1.3.

Text: font —Times New Roman, font size — 14, paragraph indent — 0.5, line spacing — 1.5, justified alignment.

The recommended manuscript size: not more than 20 pages of the specified format.

Use only standard abbreviations. Avoid abbreviations in the title and abstract. The full term for which an abbreviation stands should precede its first use in the text unless it is a standard unit of measurement.

All pages of the manuscript should be numbered.

The templates for the manuscript typesetting are presented on site:

<http://www.ipiran.ru/publication/collected/template.doc>

8. Articles should enclose data both in **Russian and English**:

- title;
- author's name and surname;
- affiliation — organization, its address with ZIP code, city, country, and official e-mail address;
- data on authors according to the format (see site):  
[http://www.ipiran.ru/journal/collected/2012\\_22\\_02\\_rus/authors.asp](http://www.ipiran.ru/journal/collected/2012_22_02_rus/authors.asp) and  
[http://www.ipiran.ru/journal/collected/2012\\_22\\_02\\_eng/authors.asp](http://www.ipiran.ru/journal/collected/2012_22_02_eng/authors.asp);
- abstract (not less than 100 words) both in Russian and in English. Abstract is a short summary of the article that can be published separately. The abstract is the main source of information on the article and it could be included in leading information systems and data bases. The abstract in English has to be an original text and should not be an exact translation of the Russian one. Good English is required. In abstracts, avoid references and formulae.
- Indexing is performed on the basis of keywords. The use of keywords from the internationally accepted thematic Thesauri is recommended.

Important! Keywords must not be sentences.

9. References. Russian references have to be presented both in English translation and in Latin transliteration.

Please take into account the following examples of Russian references appearance:

**Article in journal:**

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

**Journal article in electronic format:**

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

**Conference proceedings:**

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma "Novye resursosberegayushchie tekhnologii nedropol'zovaniya i povysheniya neftegazootdachi"* [6th Symposium (International) "New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact" Proceedings]. Moscow. 267–272.

**Books and other monographs:**

Lindorf, L. S., and L. G. Mamikonants, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

**Dissertation and Thesis:**

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticeskogo slovarya informacionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

**State standards and patents:**

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoj samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

References in Latin transcription are presented in the original language.

References in the text are numbered according to the order of their first appearance; the number is placed in square brackets. All items from the reference list should be cited.

10. Manuscripts and additional materials are not returned to Authors by the Editorial Board.
11. Submissions of files by e-mail must include:
  - the journal title and author's name in the "Subject" field;
  - an article and additional materials have to be attached using the "attach" function;
  - an electronic version of the article should contain the file with the text and a separate file with figures.
12. "System and Means of Informatics" journal is not a profit publication. There are no charges for the authors as well as there are no royalties.

**Editorial Board address:**

IPI RAN, 44, block 2, Vavilov Str., Moscow 119333, Russia

Ph.: +7 (499)135 86 92, Fax: +7 (495)930 45 05

e-mail: [rust@ipiran.ru](mailto:rust@ipiran.ru) (to Prof. Rustem Seyful-Mulyukov)

[http://www.ipiran.ru/english/journal\\_systems.asp](http://www.ipiran.ru/english/journal_systems.asp)

# **SYSTEMS AND MEANS OF INFORMATICS (СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ)**

**SCIENTIFIC JOURNAL**

**Volume 24 No.3 Year 2014**

**Editor-in-Chief and Chair of Editorial Council**  
Academician I. A. Sokolov

## **I N T H I S I S S U E:**

MATHEMATICAL SOFTWARE TOOLS FOR ANALYTICAL MODELING IN STOCHASTIC SYSTEMS WITH COMPLEX NONLINEARITIES <i>I. N. Sinitsyn, V. I. Sinitsyn, I. V. Sergeev, E. R. Korepanov, V. V. Belousov, and V. S. Shorgin</i>	4
SECURE ARCHITECTURE OF DISTRIBUTED SYSTEMS <i>A. A. Grusho, N. A. Grusho, E. E. Timonina, and S. Ya. Shorgin</i>	18
COVERT CHANNELS ON THE WIRELESS NETWORKS OF THE STANDARD 802.11 <i>N. A. Grusho</i>	32
FUSED MULTIPLY-ADD: METHODOLOGICAL ASPECTS <i>I. Sokolov, Y. Stepchenkov, S. Bobkov, Y. Rogdestvenski, and Y. Diachenko</i>	44
SELF-TIMED FUSED MULTIPLY-ADD UNIT: PRACTICAL IMPLEMENTATION <i>Y. Stepchenkov, Y. Diachenko, Y. Rogdestvenski, N. Morozov, D. Stepchenkov, A. Rogdestvenskene, and A. Surkov</i>	63
MODERN TRENDS IN EVOLUTION OF INTEGRATED NETWORK PROCESSOR ARCHITECTURES <i>V. Egorov</i>	78
TWO-PARAMETRIC ANALYSIS OF MAGNETIC-RESONANCE IMAGES BY THE MAXIMUM LIKELIHOOD TECHNIQUE IN COMPARISON WITH THE ONE-PARAMETRIC APPROXIMATION <i>T. V. Yakovleva and N. S. Kulberg</i>	92
PROPERTIES OF WINDOW DISPERSION OF MYOGRAM AS A STOCHASTIC PROCESS <i>M. Sh. Khaziakhmetov</i>	110
EXPERIMENTAL RESEARCH OF LABORATORY PROTOTYPE OF AN ARTIFICIAL HETEROGENEOUS SYSTEM FOR DIAGNOSIS OF ARTERIAL HYPERTENSION <i>I. A. Kirikov, A. V. Kolesnikov, and S. B. Rumovskaya</i>	121